R E P O R T R E S U M E S

ED 003 897 48

THE ROLE OF TRANSCRIPTION IN THE LEARNING OF THE ORTHOGRAPHIC REPRESENTATIONS OF RUSSIAN SOUNDS.

BY- CROTHERS, EDWARD AND OTHERS
STANFORD UNIV., CALIF., INST.FOR MATH.STUD.SOC.SCI

REPORT NUMBER NDEA-VI-92-4 PUB DATE 17 JUN 63
REPORT NUMBER TR-56
EDRS PRICE MF-$0.18 HC-$2.72 68P.

DESCRIPTORS- *RUSSIAN, *LANGUAGE INSTRUCTION, *LEARNING PROCESSES, *PHONEME, *ORTHOGRAPHIC SYMBOLS, EXPERIMENTAL PROGRAMS, MULTISENSORY LEARNING, COLLEGE INSTRUCTION, STANFORD, CALIFORNIA

THE USE OF PHONEMIC TRANSCRIPTION IN SECOND LANGUAGE LEARNING WAS INVESTIGATED, SPECIFICALLY AS TRANSCRIPTION FACILITATES AUDITORY-VISUAL RECOGNITION OF RUSSIAN SOUNDS AND SCRIPT. PHONEMIC TRANSCRIPTION SERVED IN THIS STUDY AS THE MEDIATING AGENT BETWEEN RUSSIAN SOUNDS AND CYRILLIC SCRIPT. THE LEARNING PROCESS WAS SET UP ON A MEDIATION CHAIN OF THREE ELEMENTS IN THE ORDER OF (1) AUDITORY STIMULUS OF RUSSIAN WORDS AND PHRASES OVER A TAPE RECORDER, (2) VISUAL STIMULUS OF THE SAME WORDS AND PHRASES, USING ROMAN SCRIPT FOR PHONEMIC TRANSCRIPTION, AND (3) VISUAL STIMULUS OF THE SAME WORDS AND PHRASES AGAIN, USING CYRILLIC CHARACTERS. TO EVALUATE THE EFFICIENCY OF THE ROMAN TRANSCRIPTIONS IN THE LEARNING PROCESS, TWO EXPERIMENTS WERE RUN--(1) ONE GROUP OF SUBJECTS WAS GIVEN TRAINING LISTS WITH THE TRANSCRIPTION PRINTED ABOVE EACH WORD, AND A SECOND GROUP LEARNED WITHOUT TRANSCRIPTION. AFTER FOUR EXPERIMENTAL SESSIONS, BOTH GROUPS WERE TESTED WITHOUT TRANSCRIPTION, (2) TRANSCRIPTION WAS UNDEREMPHASIZED IN THIS EXPERIMENT BY PRESENTING IT IN ONLY ONE-HALF OF THE TRIALS. COLLEGE-LEVEL STUDENTS WERE THE SUBJECTS OF THESE EXPERIMENTS. IN BOTH EXPERIMENTS, PHONEMIC TRANSCRIPTION APPEARED TO BE OF LIMITED VALUE IN FACILITATING THE LEARNING OF RUSSIAN. (JH)

ED0038897

92-4

THE ROLE OF TRANSCRIPTION IN THE LEARNING OF THE ORTHOGRAPHIC REPRESENTATIONS OF RUSSIAN SOUNDS

No. 4 - -

BY
EDWARD CROTHERS, RUTH WEIR AND PATRICIA PALMER

TECHNICAL REPORT NO. 56
JUNE 17, 1963

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Applied Mathematics and Statistics Laboratories
STANFORD UNIVERSITY
Stanford, California

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

THE ROLE OF TRANSCRIPTION IN THE LEARNING OF THE ORTHOGRAPHIC REPRESENTATIONS OF RUSSIAN SOUNDS

by

Edward Crothers, Ruth Weir and Patricia Palmer

TECHNICAL REPORT NO. 56

June 17, 1963

PSYCHOLOGY SERIES

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Applied Mathematics and Statistics Laboratories
STANFORD UNIVERSITY
Stanford, California

TABLE OF CONTENTS

	EXPERIMENT I	PAGE
1.	INTRODUCTION	1
2.	METHOD	4
3.	EMPIRICAL RESULTS	14
4.	LINGUISTIC INTERPRETATION OF RESULTS	20
	EXPERIMENT II	
5.	INTRODUCTION	27
6.	METHOD	28
7.	EMPIRICAL RESULTS	31
8.	REFERENCES	36
	APPENDICES A, B	37

1. INTRODUCTION

In this paper we report the third in a series of research studies on second language acquisition--aspects of the Russian linguistic system as learned by American students.[1] The first report dealt with learning to discriminate between auditorily presented consonant phonemes in CV pairs. The subject was to say whether the two members of each pair were the same or different in sound. The second report investigated vowel matching in a multiple choice situation. A monosyllabic Russian word was presented auditorily, followed by four response alternatives, also monosyllables. On each trial the subject indicated which alternative had the same vowel sound as the stimulus. The data from these two studies were subjected to quantitative and linguistic analyses (Suppes, Crothers, and Weir, 1962a, b). Both experiments dealt only with the learning of phonological features of the language, since no orthographic representations were employed.

The aim of the present study was to extend our investigation to include the learning of auditory-visual associations. On each trial, a Russian word was played over a tape recorder, and the subject saw three printed Russian words in the Cyrillic alphabet on a dittoed sheet. His task was to circle the alternative which had a certain feature in common with the auditory stimulus word (e.g., the same stressed vowel). The critical feature varied from list to list, and at the start of each list the subject was instructed appropriately. After responding, he saw the correct answer.

---

[1] The research reported herein was conducted pursuant to contract SAE 9514 between Stanford University and the U.S. Office of Education, and was also partially supported by a grant from the Carnegie Corporation of New York. We wish to acknowledge the assistance of Elise Belenky in preparing and conducting the experiment.

The possibility of presenting phonemic transcription as an experimental variable suggested itself. There seems to be little experimental evidence on the effectiveness of phonemic transcription in facilitating second language learning. Furthermore, since the tasks associated with second language learning are varied, the question must be raised whether phonemic transcription aids in the particular task we are presently considering, namely, auditory-visual recognition. Since most of the Cyrillic shapes were new to the subjects, it seemed reasonable to assume that recognition of auditory stimuli would be facilitated if the subject could see familiar visual representations. Thus, phonemic transcription could serve here as a mediating agent. That is, the hypothesized learning process could be: Russian (auditory) ⟷ Roman characters (visual) ⟷ Cyrillic characters (visual).

However, one can advance at least two counterarguments to the assertion that phonemic transcription is beneficial. First, the second link in the above hypothetical mediation chain, that between the English character and the corresponding Cyrillic character, might be difficult to learn. Second, it is possible that the presence of transcription encourages over-reliance on the transcription at the expense of learning the auditory-Cyrillic association. To evaluate the efficiency of transcription, two experiments were run. In Experiment I, one group of subjects was given training lists with the transcription printed above each word, and a second group learned without transcription. During the fifth experimental session, both groups were tested without transcription. In Experiment II, we sought to avoid overemphasizing the transcription by presenting it on only half of the trials.

Another question raised in Experiment I was whether additional auditory exposures would facilitate learning. Therefore, the transcription and no-transcription groups were each divided into two subgroups, one with and one without, auditory presentations of the choice alternatives.

A further aim of Experiment I was to gather data on the relative difficulty of phonemic discriminations in varying shapes of words and at different places in the word. Thus, the stimuli were of CVC, CV́CV, CVCV́, etc. shapes. Hence, different tasks were presented. For example, the task might be recognition of the final consonant, matching of the stressed or unstressed vowel, or matching of entire words, depending on the particular experimental list. Each subject received a sequence of lists which differed in the type of stimulus involved. Because of this periodic introduction of new material, it was not possible to collect adequate data on any one task to permit detailed testing of mathematical learning models.

Experiment II was limited to the task used in the final lists of Experiment I, word recognition in sentences. In addition to continuing the investigation of transcription, a second question asked was whether there would be any sustained improvement in performance if a second learning session were given one day after the first session, compared to a second session two weeks after the first session. One would not necessarily expect to find a permanent difference, since relearning during the second session could abolish any differences which appeared at the start of that session.

Briefly, then, both Experiments I and II were 2 x 2 factorial designs. Presence vs. absence of phonemic transcription was one variable in both experiments. The second variable in Experiment I was presence vs. absence of auditory presentation of the choice alternatives. In Experiment II the second variable was short vs. long inter-session interval.

2. EXPERIMENT I[2]

Method

Subjects and experimental design. Seventy-eight students enrolled in the summer session at Stanford University served as subjects.[3] All were public school teachers. One group consisted of elementary school teachers enrolled in Spanish classes which did not have foreign language prerequisites. The other group consisted of high school teachers of either French or Spanish whose objectives at the university were to improve their language skill and to acquaint themselves with new teaching techniques and applications of linguistics to second language teaching. More than half of the participants in both programs worked toward advanced degrees. About two thirds of the subjects were women, ranging in age from about 25 to 55. Participation in the experiment was part of their regular summer session work load.

Four groups of subjects heard a Russian stimulus word or sentence and saw the Cyrillic choice alternatives typed in capital letters. The groups were differentiated according to whether supplementary phonemic transcription of the alternatives and/or auditory presentation of the alternatives were provided. That is, the presence vs. absence of (a) phonemic transcription and (b) auditory presentation of choice alternatives, were varied orthogonally, thus:

Phonemic transcription used?	Choice alternatives presented auditorily? no	yes
no	Group $P\bar{A}$	Group $\bar{P}A$
yes	Group PA	Group $\bar{P}\bar{A}$

---

[2] The first two authors are responsible for Experiment I, and the first and third authors for Experiment II.

[3] Data from only 64 subjects were included in the analysis. The remainder were rejected because of absence from experimental sessions or because they had prior training in Slavic languages.

For example, Group PA subjects saw the material in both Cyrillic orthography and phonemic transcription, and heard both the stimulus and the alternatives. Groups $\overline{PA}$, $\bar{P}A$, $P\bar{A}$, and PA contained 13, 17, 21, and 13 subjects, respectively who completed the experiment and had no Slavic language background. Failures to complete the experiment were nearly uniformly divided among the groups, but more subjects from Groups $\overline{PA}$ and PA had received some prior Slavic language training and hence those groups are smaller than the other two. Subjects were randomly assigned to groups so that each of the four language classes contributed about one-fourth to each of the above groups.

Recordings. The tape recordings for the experiment were made in a room about 10' by 10', with a small rug acting as a sound-absorber. The recorder used was an Ampex model 601 with an Electrovoice 623 microphone and Ampex Professional Recording Tape of 1 1/2 Mil. acetate at 7.5 inches per second. The native Russian speaker who recorded for our previous experiments (Suppes et al, 1962) also recorded the present material.

Play-back apparatus. The language listening room in Stanford's main library was the play-back locale. The installation is intended for listening only. The area consists of a large carpeted room, containing a glass enclosed cubicle in one corner. Ten Ampex 610 or 612 half-track reproducing machines are located there, the output of which is fed through a 5-watt amplifier to a patch panel. In the outer area, there are fifteen tables of varying sizes, each one of which is connected by cable to the patch panel. Along the edge of each table are a number of phone jacks, one serving each student position. Through the patch panel, any one of the source machines can be connected to any or all tables. Trimm (Magnetic, 4 megohms) headphones were used.

Materials. Tape recordings were prepared of the lists of experimental items and of the instructions for each list. (By an item, we mean a stimulus word and its associated set of three choice alternatives.) Two recordings were made of each list, one for Groups $\overline{PA}$ and $P\bar{A}$, and another for Groups $\bar{P}A$ and PA. For the $\overline{PA}$ and $P\bar{A}$ groups, only the stimuli were recorded; for the other groups, both the stimuli and the choice alternatives were put on tape. Only the stimuli were taped for the test. The visual material was presented on 8 x 11 in. mimeographed sheets. On a sheet, each row corresponded to a single taped item and consisted of four columns of one word (or one CV pair) each. The entries in Columns 2, 3, and 4 were the choice alternatives. Each Russian word was typed in Cyrillic capital letters. For the $P\bar{A}$ and PA groups, the phonemic transcription was typed in lower case above the Cyrillic. In the lists involving CVCV items, the stress mark was typed with both the Cyrillic and phonemic forms. Column 1 of each item contained the correct answer to the preceding item. Each subject had a blotter with which to cover the correct answer to the item being presented. The sheets for Test I were typed in Cyrillic only.

Characteristics of the lists: phonological and graphemic features taught and criteria applied to the wrong alternatives offered. The three choices available in each item consisted of one correct and two incorrect alternatives. All alternatives had the same form (CVC, etc.) as the stimulus. Now we shall describe the individual lists. Table 1 gives the overall list structure. A sample of one item from each list appears in Appendix A.

Insert Table 1 about here

Table 1

Schedule of List Presentation and Composition of Lists

Session	List	Run[1]	Number of Items Per List	Structure[2]	Feature that is Identical in Stimulus and Correct Response Alternative
1	1	1	60	CV-CV pairs	CV-CV pair
	2	1	60	" "	" "
	3	1	60	" "	" "
2	4	1-3	162	CVCV	stressed vowel
	5	1-2	50	CVCV regular	unstressed vowel
3	5	3-4	50	CVCV regular	unstressed vowel
	6	1-4	120	CVCV irregular	unstressed vowel
3	7	1-3	150	CVC, CVCC, CCVC regular	final consonant
4	8	1-2	100	CVC, CVCC, CCVC irregular	final consonant
4	Test I	---	90	items from lists 1-3, 5, 6, 7, 8	(feature determined by list from which item selected)
5	9	1	60	two word sentence form; word + CVCV or CVCV + word	CVCV word
5	10	1-2	100	same as in List 9 except that CVC, CCVC, CVCC replaced CVCV	CVC, CCVC, or CVCC word
5	11	1	50	longer sentences; critical word may have more than 2 syllables	word

[1]Each run is a different randomization of the items in the list.

[2]See the text for a more detailed description of the make-up of each list and Appendix A for examples.

Lists 1-3 consisted of CV pairs. For half the items, the stimuli (and, hence, the correct answers) may be symbolized as $C_1V_1 - C_1V_1$. Pairs of this type were called s ("same") pairs in our first report (Suppes et al., 1962a). The stimuli and the correct answers in the other half of the items were $C_1V_1 - C_2V_1$, or d ("different") pairs. In Lists 1-3 the second CV syllable of each pair was the same in the incorrect as in the correct alternatives, so the subject was required to discriminate some feature of the first CV syllable. The second syllable served merely to permit an immediate comparison with the critical syllable. In List 1, the consonant in the initial syllable of one wrong alternative was phonologically similar to that in the correct alternative. In the other wrong alternative, the initial consonant was graphemically similar to that in the correct alternative. List 1 differed from List 2 in that the former utilized the phonological contrasts of voiced vs. voiceless consonants, while the latter presented contrasts of different manners of articulation, e.g., stop vs. fricative, fricative vs. affricate. In most cases the consonant phonemes were unpalatalized. List 3 differed in that the phonological discriminations were between unpalatalized consonants in the first CV syllable; all wrong alternatives contained the palatalized counterpart of the correct unpalatalized consonant phoneme.

Table 2 summarizes the item structure for Lists 1-3. Table 3 indicates which graphemes were designated similar to one another in constructing the

---

Insert Tables 2 and 3 about here

---

lists. If no graphemes could be found which were similar to that of the initial grapheme of the stimulus, the choice of the graphemically wrong alternative was arbitrary. Within each run (repetition of the same list of items) the items were presented in random order.

Between lists there was, in general, a progressively increasing order of difficulty. Within lists, we decided to explore a wide range of lexical variation, rather than to attempt to equate the items in difficulty or structural properties. Lexical items can vary on so many dimensions (to name only a few, choice of initial consonant, choice of vowels in the wrong alternatives, similarity of a phoneme in the stimulus to one in a right or wrong alternative) that it seemed wise to pursue the study of responses on this wide structural range in lieu of a detailed investigation of a few properties of unknown relative importance.

List 4 was composed of CVCV pairs. The subjects' task was to select the alternative with the same stressed vowel as that in the stimulus. The items were about evenly divided between the $C\acute{V}CV$ and the $CVC\acute{V}$ forms. Also, the correct stressed vowel appeared in the same syllable in the stimulus as in the correct answer about half the time. It seemed likely that the subject could err by correctly identifying the stressed syllable but not the particular vowel (e.g., selecting $C\acute{V}_1CV$ when the correct answer was $C\acute{V}_2CV$). Therefore, each item contained one alternative which had the stress on the correct syllable but was the wrong vowel. A type of error hypothesized to be less likely was that the subject would correctly identify a vowel but be uncertain of the locus of the stress. This might lead, for example, to $CV_1C\acute{V}$ responses (errors) to $CVC\acute{V}_1$ stimuli. To measure this tendency, each item contained one alternative which included the correct vowel, but in the unstressed form. The 54 items in List 4 included 12, 6, 12, 12, 12, having stressed vowels /a, e, i, o, u/, respectively. A different random order was used with each of the four presentations of the list.

Table 2

Phonemic Make-up of Lists 1-3

List	Contrast	Vowel in Correct Answer
1.	/t:d/	/a,e,i,o,u/
	/k:g/	/a,o/
	/p:b/	/i,o,u/
	/s:z/	/a,o/
	/f:v/	/u/
	/sh:zh/	/a,i,o/
2.	/k:x/	/a,e,i,o,u/
	/sh:šč/	/a,e,i,o,u/
	/s:ts/	/a,e,i,u/
	/ch:ts/	/a,o,u/
	/ch:tj/	/i,o,u/
3.	/m:mj,z:zj,r:rj,d:dj/	/a/
	/t:tj,d:dj,p:pj,b:bj, s:sj,z:zj,l:lj,r:rj m:mj,n:nj/	/i/
	/n:nj,r:rj,b:bj,s:sj, m:mj/	/o/
	/l:lj,m:mj,t:tj,r:rj/	/u/

Table 3

Similar Russian Graphemes

Е	Ё
Э	З
И	Н
У	Ч
Л	Д,П
Ш	Щ
Ж	Ш,Щ
Ц	Щ

List 5 also consisted of CVCV words, but now the subjects' task was to select the alternative with the same unstressed vowel as that in the stimulus. Only "regular" unstressed vowels appeared in the stimuli, that is, the unstressed o symbol and the unstressed e symbol were not included, due to the reduced Russian vowel pattern in unstressed positions which obliterates the phonemic distinction of high vs. non-high with back and front vowels. Of the List 5 stimuli, 10, 6, and 9 involved unstressed /a/, /i/, and /u/, respectively. All five vowels were used in the unstressed position in the wrong alternatives. This reduced the item difficulty, since subjects could learn that unstressed o and e were never correct, but afforded opportunity (except in Group $\overline{PA}$) for learning the discrepancy between sound and symbol for these two vowels. As in List 4, no two alternatives in an item contained the same stressed vowel, the same unstressed vowel, or the same initial consonant. One of the wrong alternatives had the same stress pattern as the stimulus but the wrong unstressed vowel. The other had as stressed vowel the unstressed stimulus vowel. As in List 4 both CV́CV and CVCV́ were used in the stimuli and alternatives. A new randomization was used with each repetition of the 25 items.

List 6 was primarily concerned with CVCV words containing the "irregular" unstressed vowel symbols o and e, which had been excluded from List 5. To prevent the correct answer from being obvious, it was necessary to include stimuli with the other unstressed vowels as well. There were 3, 10, 4, 10 and 3 items with unstressed a, e, i, o, u, respectively. The other features of List 6 were the same as in List 5. A new randomization was used with each repetition of the 30 items.

List 7 involved CVC words, in addition to shapes with consonant clusters, namely, CVCC, CCVC, CCCVC and CVCCC. The subjects' task was to select the alternative with the same final consonant as the stimulus. Forms having voiced final stop and fricative symbols were deferred to List 8, since no such voiced consonant occur word-finally in Russian. The three alternatives in a given item contained no phonemes in common with each other. However, in many cases they contained a vowel or consonant in common with the stimulus. Ten CVC items, then 20 CVCC (or CVCCC), and finally 20 CCVC (or CCCVC) were used; both plain and palatalized endings occurred. The incorrect and correct alternatives did not differ in such a systematic fashion as in Lists 1-6. For example. if the correct ending were t, the incorrect endings might be l, n, etc. in one item, and k, ch, etc. in another item. A different random presentation order was used in each repetition of the list.

As in List 7, List 8 items were also CVC, CVCC and CVCCC, and CCVC and CCCVC. However, the only forms used were those omitted from the preceding list, those words having voiced final consonant letters with corresponding voiceless phonemes. In other respect, the structures of Lists 7 and 8 were similar.

Following List 8, the test was administered. The test consisted of a sampling of items from Lists 1, 2, 3, 5, 6, 7, and 8. The subjects' task on a particular test item was the same as when that item had appeared in a preceding list. Since List 5 was represented, no items for List 4 appeared. Otherwise, ambiguity would arise regarding whether the subject was to respond on the basis of the stressed or the unstressed vowel. Test I contained 8, 8, 10, 15, 15, 15, 15, and 14 items from Lists 1-8 respectively, or a total of 100 items.

List 9 introduced two-word sentences. One of the two words was a CVCV pattern. The structure of the other word was not controlled. In half the sentences, the CVCV word preceded the other word. Now the task was to circle the alternative which was the visual representation of the CVCV member, rather than to match a specific phoneme of the stimulus. The items were presented in random order.

List 10 differed from List 9 in that the critical word was of CVC, CCVC, or CVCC form. There were 10, 20, and 20 sentences, respectively, with these types. Other features of List 10 were the same as in List 9.

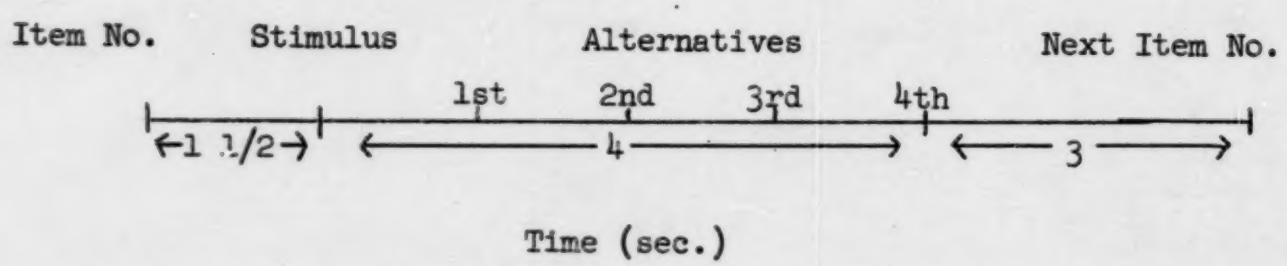
List 11 explored word recognition in longer sentences. The first 40 items were three or four word sentences classified according to the critical word, divided thus: 10 CVCV, 5 CVC "regular", 5 CVC "irregular", 5 CVCC "regular", 5 CVCC "irregular", 5 CCVC "regular", and 5 CCVC "irregular". The last ten sentences were of more than four words each. The position in the sentence of the critical word varied from one item to the next. The 50 items were presented in random order.

Time intervals. Care was taken during recording to ensure that the several runs of the same lists of items were of equal duration. Also, it was important to require that the tapes for the $\overline{PA}$ and $P\overline{A}$ groups proceeded at the same rate as those for the $\overline{P}A$ and PA groups. On the other hand, the nature

of the items and the fixed time available during a session dictated that the total presentation time vary from one list to another. This variation was achieved by appropriately specifying the number of items in each run, not by artificially increasing or decreasing the speaking rate per item. The average duration of a run in each list was:

List	Time per run (min.)	List	Time per run (min.)
1	10.5	7	6.9
2	11.2	8	8.3
3	11.6	Test I	12.0
4	7.1	9	7.5
5	7.2	10	10.0
6	4.3	11	16.1
		Test II	21.0

Procedure. Each subject was seated at a table in the language listening room and supplied with earphones and answer sheets. He listened to the stimulus member of each item and saw on his sheet the Cyrillic representation of each choice alternative. Subjects in Groups $P\bar{A}$ and PA also saw the phonemic form of the word or words. Subjects in Groups $\bar{P}A$ and PA heard the choice alternatives as well as the stimulus. The subject responded by circling an alternative. Next he moved his blotter down a line on his sheet, and looked in the left column to see if his last answer was correct. The time intervals between successive taped events were:



The procedure continued without interruption for about twelve minutes. Then there was a short pause to permit changing of the tape reels, and the items continued as before. Appropriate instructions were played before each list. The instructions for Lists 1-3 for Groups $P\bar{A}$ and $\bar{P}A$ were:

> Lists 1-3, CV Pairs. Place the blotter so that item number 1 on your answer sheet is exposed and item number 2 is covered. Now look at item number 1. It contains three pairs of syllables printed in Russian. You are going to hear one of these three pairs of syllables and your task is to decide which of the three pairs you heard. If you have the English spelling printed above the Russian, you may use this to help you in answering. Some of the Russian letters will have unfamiliar shapes, but you will still have grounds for making a choice. After you have answered, move the blotter down one line and look at the word in the first column to see if you were right. Then listen to the next pair of syllables and continue as before. Remember, each time you listen, circle an answer, move your blotter down one line, and look at the word in the first column to see if you were right. Be sure to answer, even if you are not certain.

The instructions for the other groups differed only in that they were told they would hear the alternatives. The Lists 4-11 and test instructions appear in Appendix B.

Schedules for experimental sessions. Since the students were run during the hour when they normally attended a language class, the schedule differed from one subject to another. These discrepancies were minimized by assigning each subject two sessions weekly until seven sessions had been completed. The first session occurred on a Monday or Tuesday, and the second on a Thursday. The duration of each session was 25 minutes.

3. EMPIRICAL RESULTS OF EXPERIMENT I

Two types of findings are considered. First, there are comparisons between the groups on learning the various lists. Secondly, we briefly present data on intralist learning. The more detailed analyses of the phonemic structures of hard and easy items are treated separately in Section 4. Some interpretations of the empirical results are offered here, but all discussions based on linguistic arguments are deferred until the next section. Of course, we must be cautious in making statements regarding interlist differences in difficulty. Our aim was to obtain an overall picture of identification learning involving a great many items presented in a fixed order. Hence, our interpretations are made with the reservations that a counterbalanced order of list presentation and different criteria for the selection of items could yield different results. However, it seems unlikely that the intergroup comparisons are subject to these qualifications.

Comparison of the groups. Table 4 shows the proportions of correct phoneme identifications on each of Lists 1-8 and the test, and of correct word identifications on each of Lists 9-11. Using the total correct for each subject, Mann-Whitney U tests were performed to ascertain whether the differences between the groups were significant, and the results are given in Table 5.

Insert Tables 4 and 5 about here

As the tables indicate, the groups with phonemic transcription available scored higher on the training Lists 1-8 than did the groups without transcription. The average proportion correct over Lists 1-8 was .65, .75, .66, and .75 for Groups $\overline{PA}$, $P\overline{A}$, $\overline{P}A$, and PA, respectively. The total correct on these lists was significantly higher ($P < .02$) for Group $P\overline{A}$ than for Group $\overline{PA}$. Likewise, the

Table 4

Proportions of Correct Phoneme Identifications on Lists 1-8 and Word Identifications on Lists 9-11

	List											
Group	1	2	3	4	5	6	7	8	Test	9	10	11
$\overline{PA}$	.67	.64	.70	.57	.58	.54	.71	.80	.55	.52	.61	.54
$P\overline{A}$	.79	.76	.82	.74	.65	.64	.80	.80	.44	.80	.87	.75
$\overline{P}A$	.69	.70	.68	.70	.51	.51	.69	.78	.47	.81	.85	.65
PA	.80	.74	.75	.80	.65	.62	.78	.82	.45	.89	.90	.90

	1-8	9-11
$\overline{PA}$	.65	.56
$P\overline{A}$	.75	.81
$\overline{P}A$	.66	.77
PA	.75	.83

Table 5

Results of Mann-Whitney U Tests

of Intergroup Differences[1]

	Lists 1-8			Test			Lists 9-11[2]	
Group	$\overline{PA}$	$\overline{\overline{PA}}$		$\overline{PA}$	$\overline{\overline{PA}}$		$\overline{PA}$	$\overline{\overline{PA}}$
$\overline{PA}$	60**	106	$\overline{PA}$	73*	58*	$\overline{PA}$	---	---
PA	132	42**	PA	136	100	PA	116	56*

* $P < .05$ ** $P < .02$

[1] An entry in a given row and column is the value of U between the row group and the column group. All significance levels are based on two-tailed tests. For Groups $\overline{PA}$, PA, $\overline{PA}$, and $\overline{\overline{PA}}$, N = 13, 13, 21, and 17, respectively.

[2] It was obvious by inspection of Table 4 that on Lists 9-11 the scores were significantly lower for Group $\overline{PA}$ than for the other groups.

performance of Group PA exceeded that of Group $\overline{P}A$ ($P < .02$). On the other hand, neither the difference between Groups PA and $P\overline{A}$ nor that between Groups $\overline{P}A$ and $\overline{PA}$ was significant. Hence, the effect of auditory presentation of the choice alternatives during the training lists was negligible.

On the other hand, the results of the test indicate that the superiority of Groups PA and $P\overline{A}$ was not maintained when the phonemic transcription was no longer present. In fact, there was a complete reversal of the relative performance of the groups, in that the groups who had not formerly seen the phonemic transcription did better on the test. There were significantly more correct responses for Group $\overline{PA}$ than for either of Groups $P\overline{A}$ or $\overline{P}A$ ($P < .05$ in both cases). However, the difference between Groups PA and $\overline{P}A$ was not significant, so prior training with transcription was a handicap only for the groups who had not heard the alternatives during training. Another way of stating the interaction is that prior auditory exposure to the alternatives hindered test performances when no transcription had been available formerly (Group $\overline{PA}$ vs. Group $\overline{P}A$). But the difference between Groups $P\overline{A}$ and PA was not significant. Since Group $\overline{PA}$ did best, the importance of identical conditions during training and testing seems to outweigh any transfer from prior exposure to transcription or auditory alternatives. This is not to say that the presence of the phonemic transcription did not produce any faciliatation in the learning of the correspondences between the sounds and the Cyrillic representations. Some positive transfer may have occurred, but it was outweighed by the heavy dependence on the transcription. It is possible that explicit instructions in the use of phonemic transcription would have produced more transfer, but this is open to question.

As the tables show, identification of words in sentences (Lists 9-11) also was definitely facilitated when phonemic transcription was present. The proportion correct over Lists 9-11 averaged .56, .81, .77, and .83 for Groups $\overline{PA}$, $P\bar{A}$, $\bar{P}A$, and PA, respectively. No statistical tests are required to see that Group $\overline{PA}$ did significantly poorer than the other groups in this final phase of the experiment. Phonemic transcription also was helpful when the alternatives were heard (Group PA vs. Group $\bar{P}A$, $P < .05$). It would seem that the word identification task required better memory of what had just been presented than did the phoneme identification task. That is, the subject had to retain an entire sentence, instead of just a single word, until he could decide which word of the sentence he was to identify. This observation may offer a clue as to why Group $\overline{PA}$ was so deficient on word identification; namely, the task of word identification is closer to natural language tasks than is mere phoneme identification, and in this context multiple clues are more relevant than in matching of individual phonemes. Even though this may be only a partial explanation for the poor performance by this group, we take comfort in noting that these results amplify, rather than offset, the overall differences between phonemic vs. no phonemic and between auditory vs. non-auditory alternatives.

Comparison of the lists. Subject to the reservations mentioned earlier concerning interlist transfer, we can make some general statements about the relative difficulty of the various lists, given our order of list presentation. Of course, these remarks are intended to be more tentative than those concerning the intergroup differences. For convenience in comparing lists, the data of Table 4 have been displayed as a "learning curve" for each experimental group (Figure 1).

Insert Figure 1 about here

The order of presentation of the various Lists 1-8 (but not the within-list presentation order) was designed to constitute a progression from easier to more difficult identifications. It was suggested that the increase in difficulty would be counteracted by the emergence of a "learning set". That is, some positive transfer was considered plausible, either via mastery of general orientation techniques (rapidly learning "what to listen for") or of specific phoneme identifications. Nevertheless, the evidence in our report on vowel identification (Suppes et al, 1962b) is that marked positive transfer does not occur under certain conditions. In that study, the group of subjects who had progressively more difficult discriminations did progressively poorer. One conclusion offered for this trend was that it is unreasonable to expect positive transfer in a task of this magnitude unless the original training is continued over a longer period of time than is customary in psychological experiments. Also, since different contexts produce allophonic variations within each phoneme, the grouping of allophones into a single phoneme is an added task for the subject in learning such progressively ordered material.

That conclusion seems to have been upheld in the present experiment, at least insofar as "transfer" is considered to be improvement over lists, independently of control groups. Not only was there an absence of improvement, but the curves in Figure 1 are conspicuously non-monotonic. Identification of regular unstressed vowels (List 5) and irregular unstressed vowels (List 6) proved to be the most difficult phases of the first part of the experiment.

For all groups, there was a sharp drop in the proportion correct on the test, as contrasted with the proportions on the eight training lists. The proportions correct on the test items which had been selected from the training lists were .42, .50, .65, .41, .37, .44, and .57 for the items from Lists 1, 2, 3, 5, 6, 7, and 8, respectively. Recall that the expected proportion correct by guessing was .33. The high frequency of errors obviously corroborates the claim that we were introducing new material at a more rapid rate than it was being learned. Specifically, we attribute the poor test performance to: (a) forgetting between the training list and the test; (b) removal of phonemic transcription and of auditory presentation of the alternatives; (c) lack of feedback on the test as to whether the response was correct; (d) greater heterogeneity of items on the test. The last point can be clarified by noting that within a training list the subject could become "set" to listen for a specific component of the stimulus. However, on the test he had to shift his set from one item to the next, as well as to remember what his task was for CVC, CVCV, etc..

Intralist learning. Although we were chiefly interested in interlist and intergroup comparisons, we also plotted the course of learning within the individual lists. For illustrative purposes, the learning curves for Lists 2, 4, 8, and 10 are shown in Figure 2. These curves exhibited no important

---

Insert Figure 2 about here

---

differences from those on the odd-numbered lists. The graphs represent the proportions correct on successive blocks of 20 trials (except in List 4, where the block size is 18). The first overall point we wish to make is that

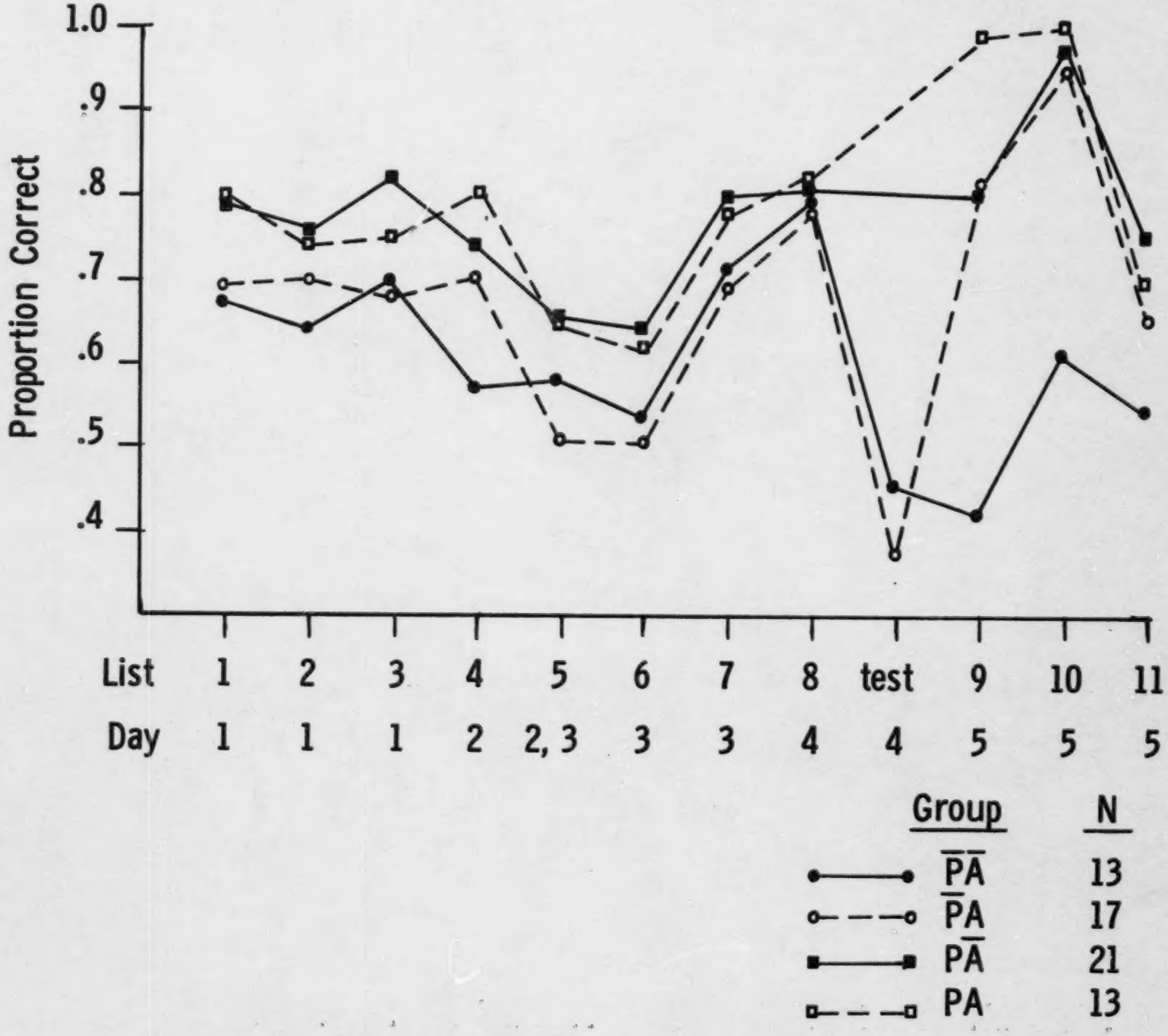


Figure 1. Proportion correct on each list in Experiment I.

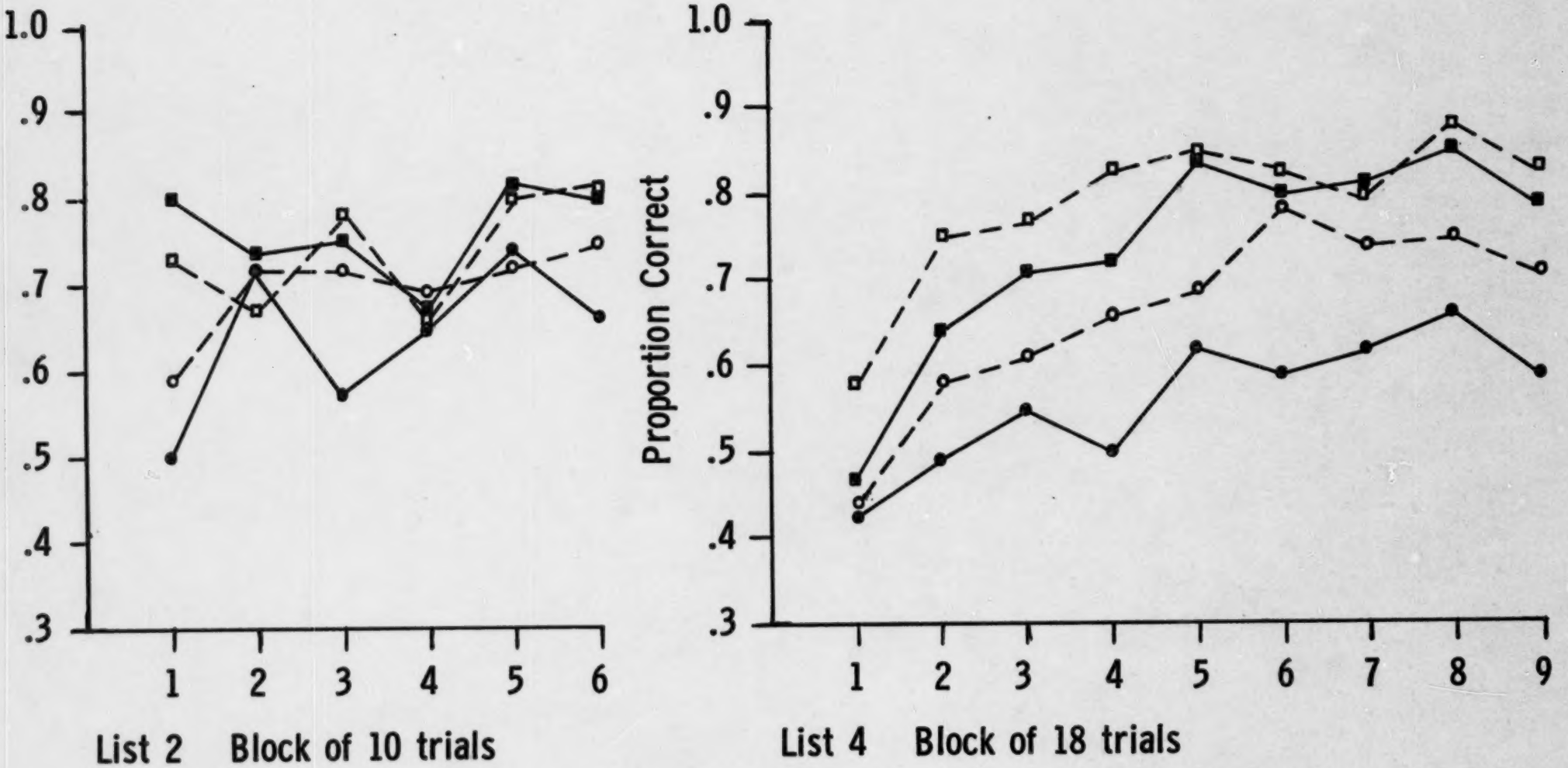


Figure 2 a. Intralist learning as illustrated by Lists 2 and 4.

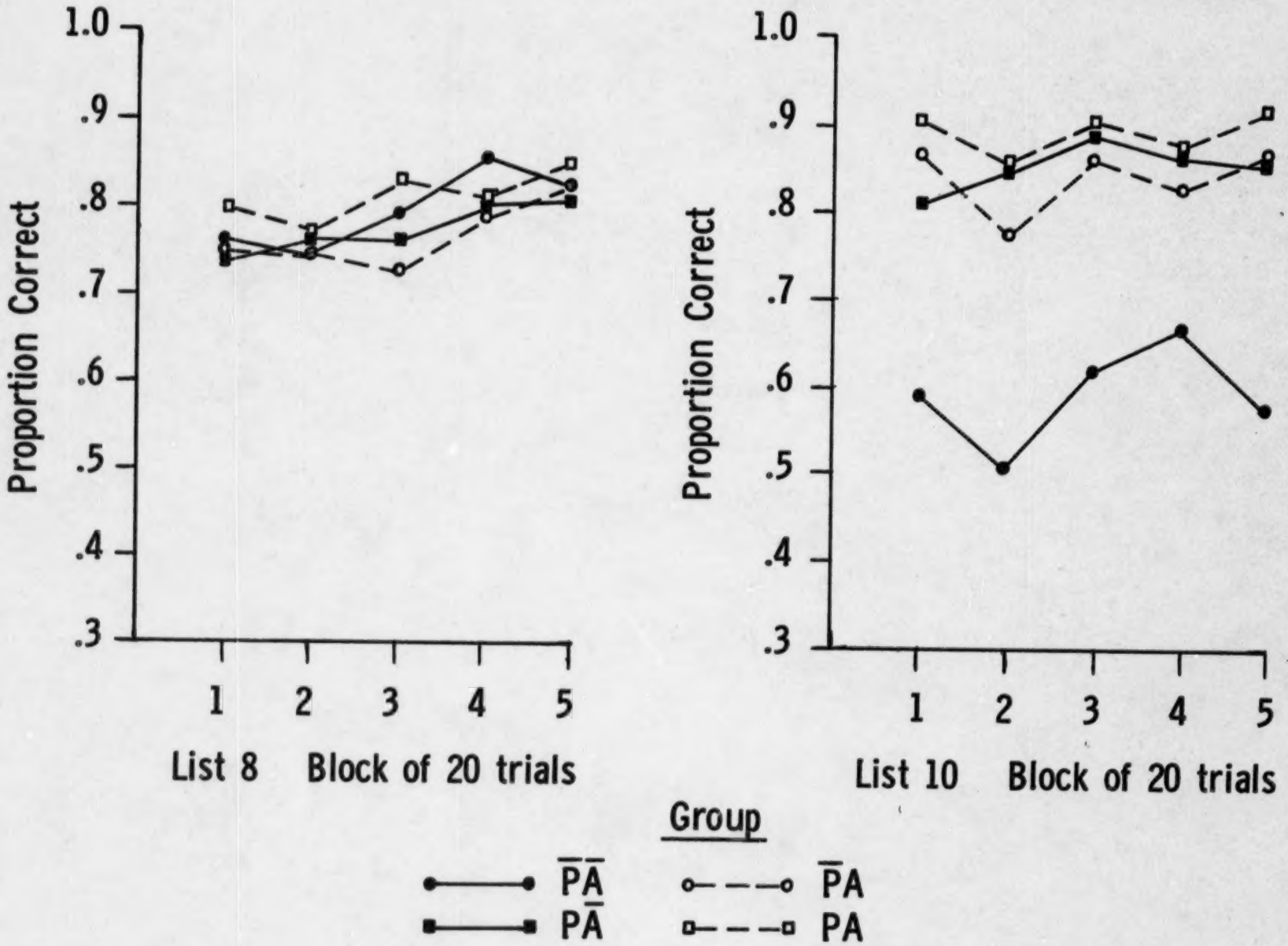


Figure 2 b. Intralist learning as illustrated by Lists 8 and 10.

intralist learning did occur, and was somewhat more prominent on the earlier lists. In fact, the rate of improvement was remarkably low on Lists 6 and 8, and was negligible on List 10. A second point may be established by referring especially to the list with the greatest number of blocks of trials (List 4). If one were to estimate the asymptotic proportion correct by visually inspecting the group curves for this list, the conclusion would be that the asymptote should be less than one.

The conclusion regarding the learning rate and that concerning the asymptote appear to have some general pedagogic implications for the effectiveness of our method of teaching phoneme and word recognition. Although the method is suitable for assessing the difficulty of a wide variety of items, it leaves much to be desired from the standpoint of producing learning. Of course, we could undertake a detailed post hoc assessment of each experimental variable which was equated in all groups in an effort to explain the slowness of learning. However, that would be somewhat beside the point for our present purpose. We maintain that a procedural change which lowers to general level of item difficulty leaves the order of relative item difficulty relatively intact (Suppes et al, 1962b).

4. LINGUISTIC INTERPRETATION OF RESULTS OF EXPERIMENT I

Since each list was constructed on the basis of linguistic categories, it seems best to present our suggested interpretations following the order of list presentation.

The first three lists, of same or minimally different pairs of CV syllables were roughly arranged according to the feature of voiceless vs. voiced, stop vs. fricative vs. affricate, and plain vs. palatalized. This is somewhat analogous to the procedure reported in our first paper (Suppes et al, 1962a), where the subjects judged whether two auditorily presented syllables of a given pair were the same or different. However, in the present experiment the subjects' task was to recognize the orthographic form of the auditory stimulus in a multiple choice task, where the wrong alternatives were either phonemically or graphemically determined. The proportions correct on List 1, for items classified according to consonant contrast, were:

Ф:В (f:v)	.80	Ш:Ж (sh:zh)	.79	С:З (s:z)	.79
Т:Д (t:d)	.76	П:Б (p:b)	.57		

The most difficult identification by far, for all experimental groups, is that of /p/ and /b/. We had obtained similar results in a previous experiment described in the above-mentioned report without being able to suggest a plausible explanation, since structurally and impressionistically the bilabial stops do not differ appreciably from the dento-alveolar or velar ones. The effect of the vowel following the consonant to be identified seemed of little relevance to the /p:b/ difficulty except perhaps that no /a/ follows the "difficult" bilabials, and half of them are followed by /i/, which we had

clearly identified as the most difficult vowel phoneme to match (Suppes et al., 1962b). However, one of the least difficult items also contains /i/, and /o/ and /u/ seem to be about evenly divided between most and least difficult items.

We therefore hypothesized that a spectographic analysis of some of the bilabial syllable pairs might yield suggestions as to why these were more erroneously matched. (See also Reeds and Wang, 1961).

To do this, a Kay Electric Sonagraph Model R was used. The stimuli were reproduced on an Ampex Model 351 full track tape recorder, using appropriate impedance matching. For each syllable pair of List 1, which involved a stop, a conventional Sonagram was made using a 300 cps bandwidth and an effective frequency range of from 85 to 4,800 cps.[4]

A comparison of the several Sonagrams thus obtained for the Russian speaker used in the experiment yielded no consistently applicable explanation of the judgmental results described. Since the listeners used were native speakers of American English, it was decided that a similar acoustic analysis of stop sounds produced by a female speaker of American English might be helpful. One speaker was asked to record the following syllable pairs:

---

[4] We wish to acknowledge gratefully the help of Mr. William Castle of the Stanford University Speech Pathology and Audiology Laboratory staff in preparing the Sonagrams.

[pi pi]	[ti ti]	[ki ki]
[pu pu]	[tu tu]	[ku ku]
[po po]	[to to]	[ko ko]
[pa pa]	[te te]	[ke ke]
[pɛ pɛ]	[tɛ tɛ]	[kɛ kɛ]
[bi bi]	[di di]	[gi gi]
[bu bu]	[du du]	[gu gu]
[bo bo]	[do do]	[go go]
[ba ba]	[da da]	[ga ga]
[bɛ bɛ]	[dɛ dɛ]	[gɛ gɛ] [5]

These syllable pairs were then subjected to Sonagraphic analysis. The resulting Sonagrams were compared to those made for the Russian speaker. The results of this comparison are summarized in Table 6, and are illustrated

---

Insert Table 6 about here

---

in the series of Sonagrams in Figures 3a -3f. The table shows that for the American speaker the bilabial stops showed no burst element, whereas the dental and velar stops did show it. On the other hand, all of the Russian stops showed a burst element. Such evidence suggests that the burst portion may not be a relevant cue for the recognition of American bilabial stops. It seems reasonable to assume that American listeners are not accustomed to the presence of such a cue and are confused by it when asked to listen for bilabial stops in Russian.

The acoustic results with regard to the aspiration element suggest that in both American English and Russian, voiced stops are generally distinguished from voiceless ones by the relative amount of aspiration which follows the burst portion. When this distinction holds the aspiration is

---

[5] International Phonetic Alphabet.

Table 6

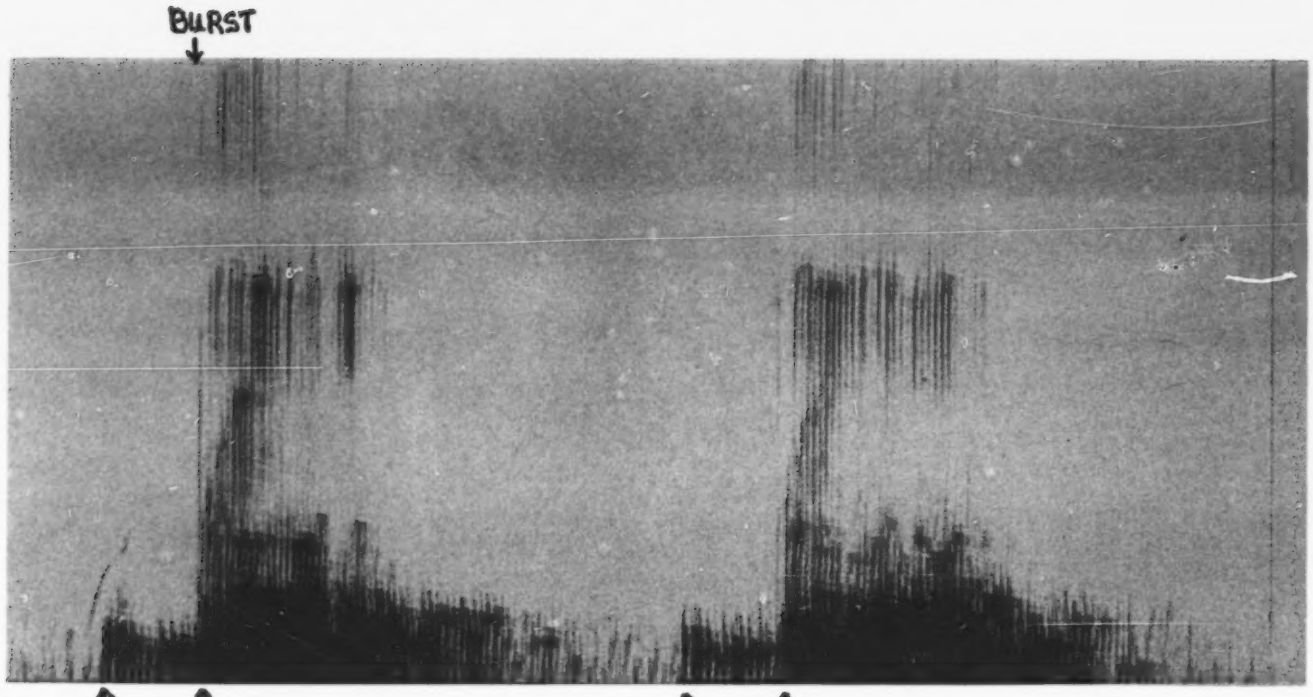
Presence and absence of voice bar, burst, and aspiration with Russian and American English bilabial, dento-alveolar, and velar stops.

LEGEND: X - Present; O - Absent		RUSSIAN		AMERICAN	
		VOICED	UNVOICED	VOICED	UNVOICED
VOICE BAR	BILABIAL	X	O	O (9 out of 10)	O
	DENTAL	X	O	O (8 out of 10)	O
	VELAR	X	O	O (7 out of 10)	O
BURST	BILABIAL	X	X	O	O
	DENTAL	X	X	X	X
	VELAR	X	X	X	X
ASPIRATION	BILABIAL	O	O	O	X (long)
	DENTAL	O	X (short)	X (short)	X (long)
	VELAR	O	X (short)	X (short)	X (long)

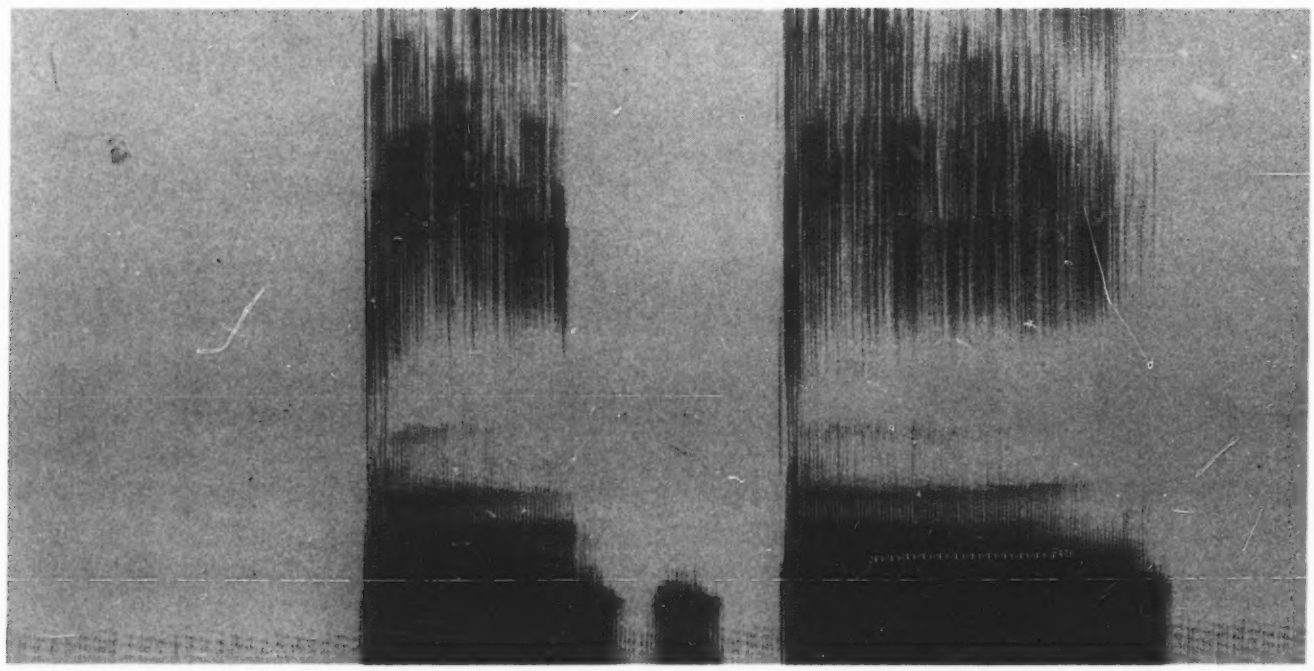
Figures 3a - 3f

Sonagrams of Russian and American English
monosyllables consisting of stops plus vowel.

VOICED BILABIALS

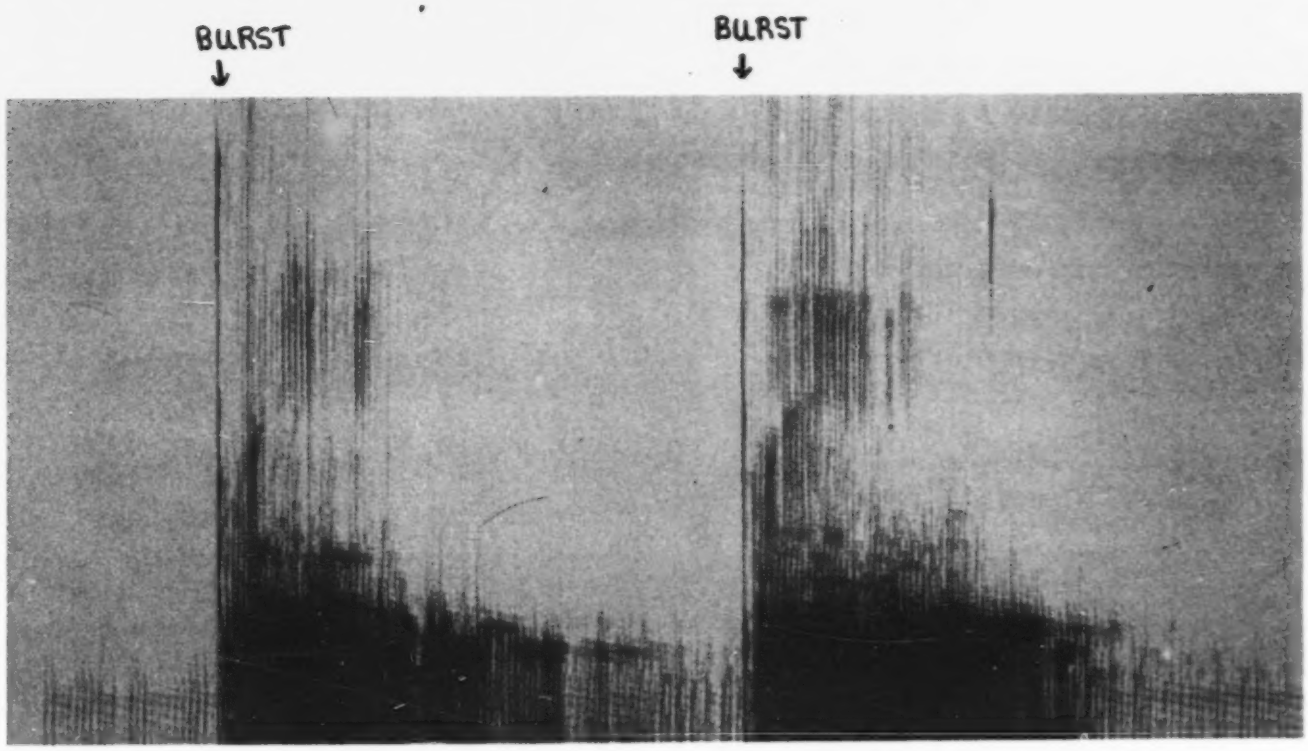


Bb Bb

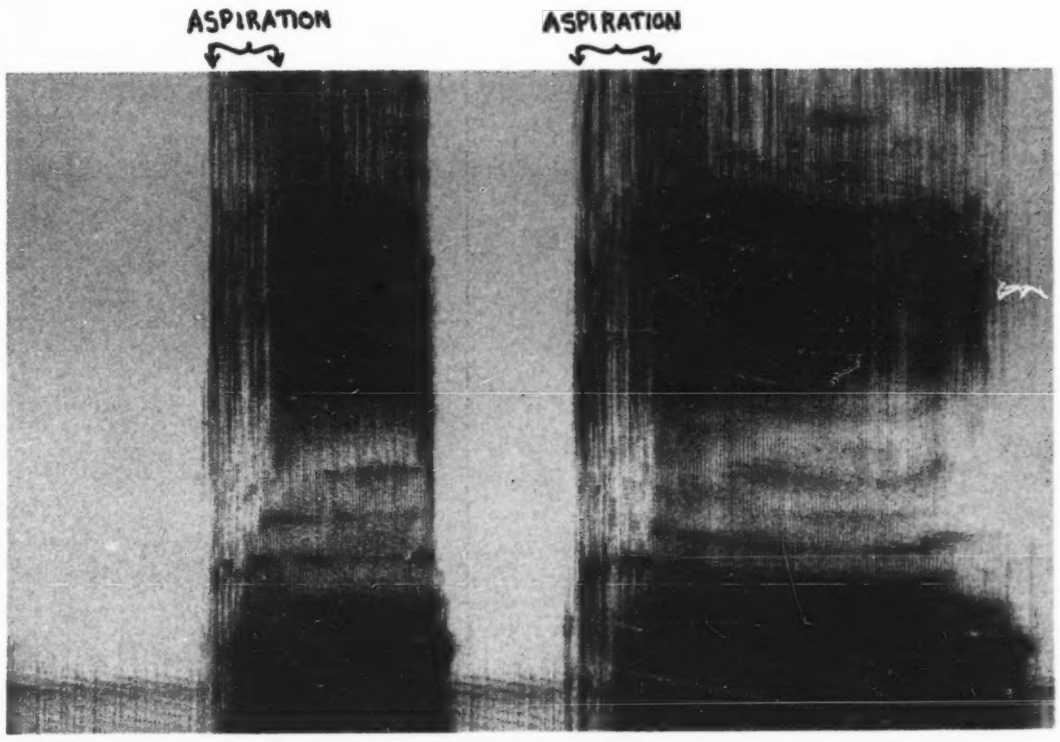


[bi] [bi]

UNVOICED BILABIALS

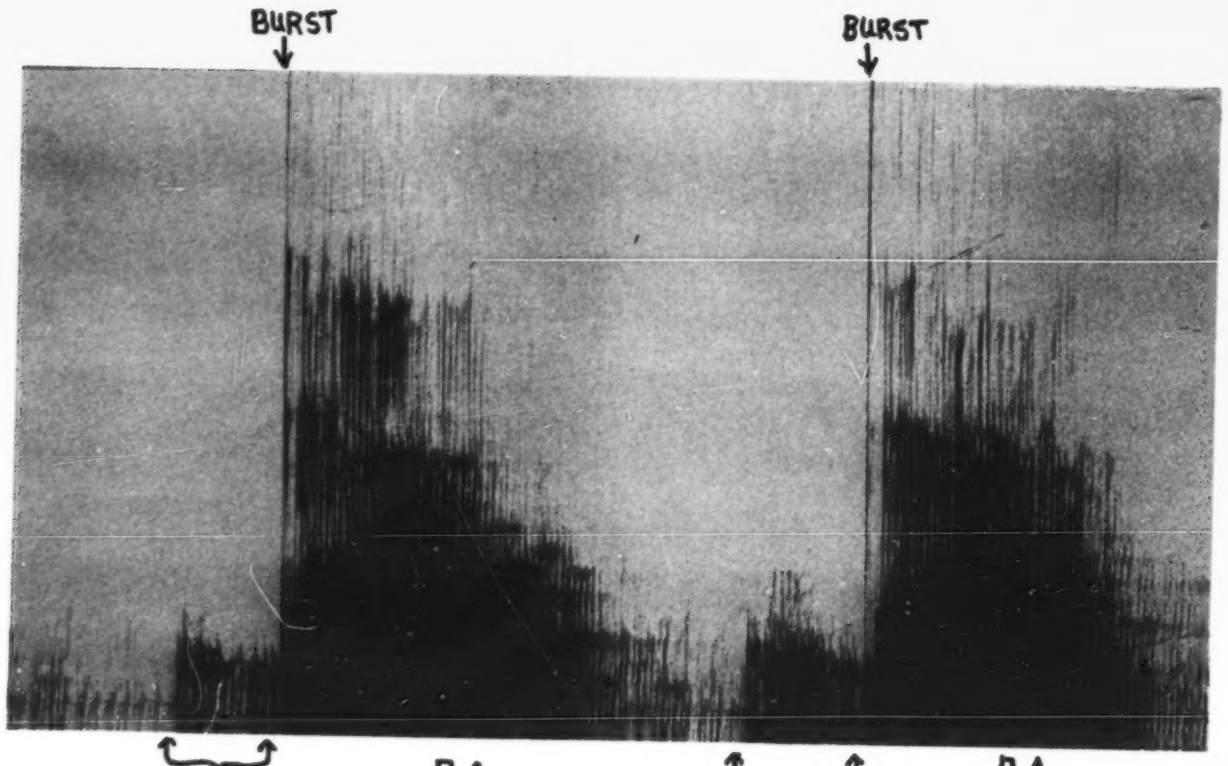


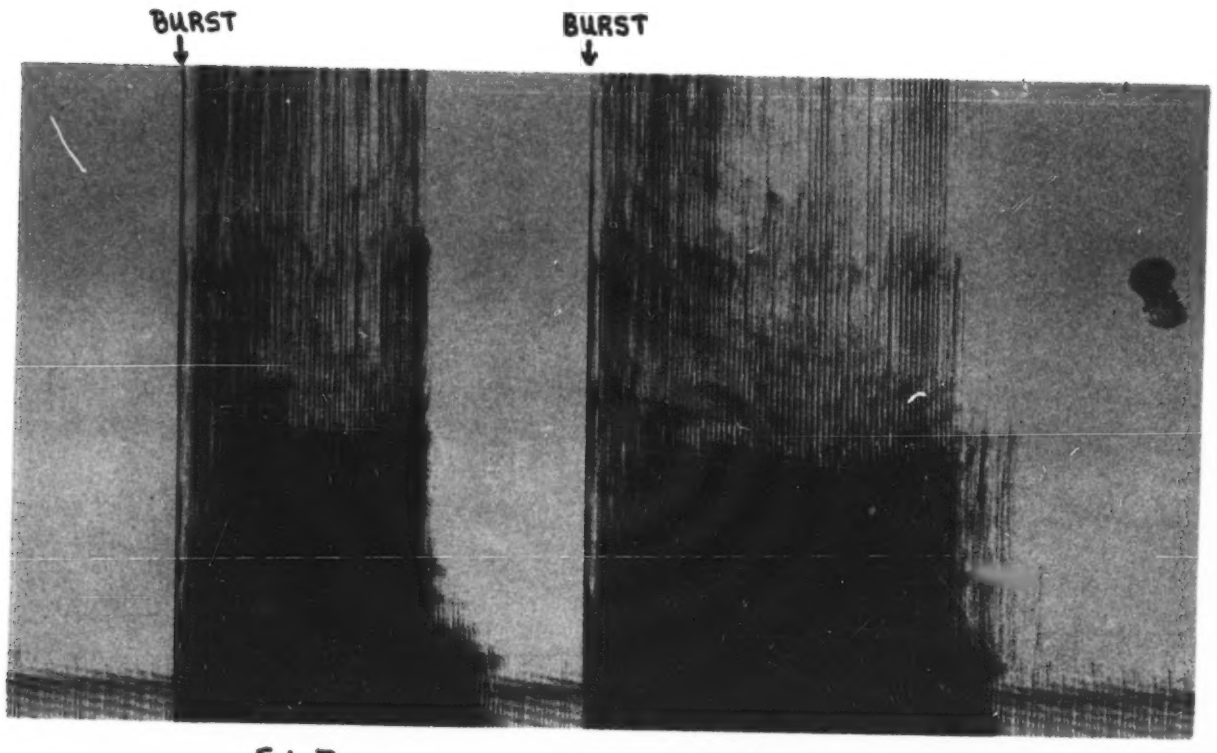
ПЫ ПЫ



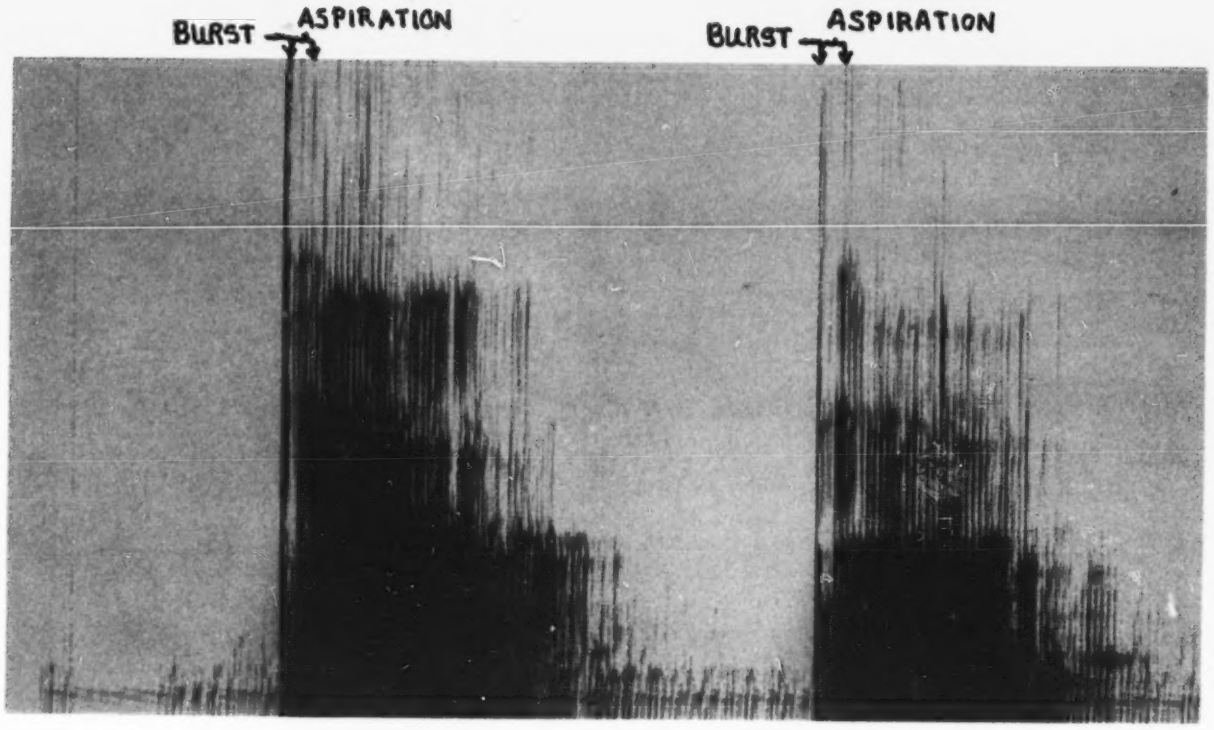
[pi] [pi]

VOICED DENTALS

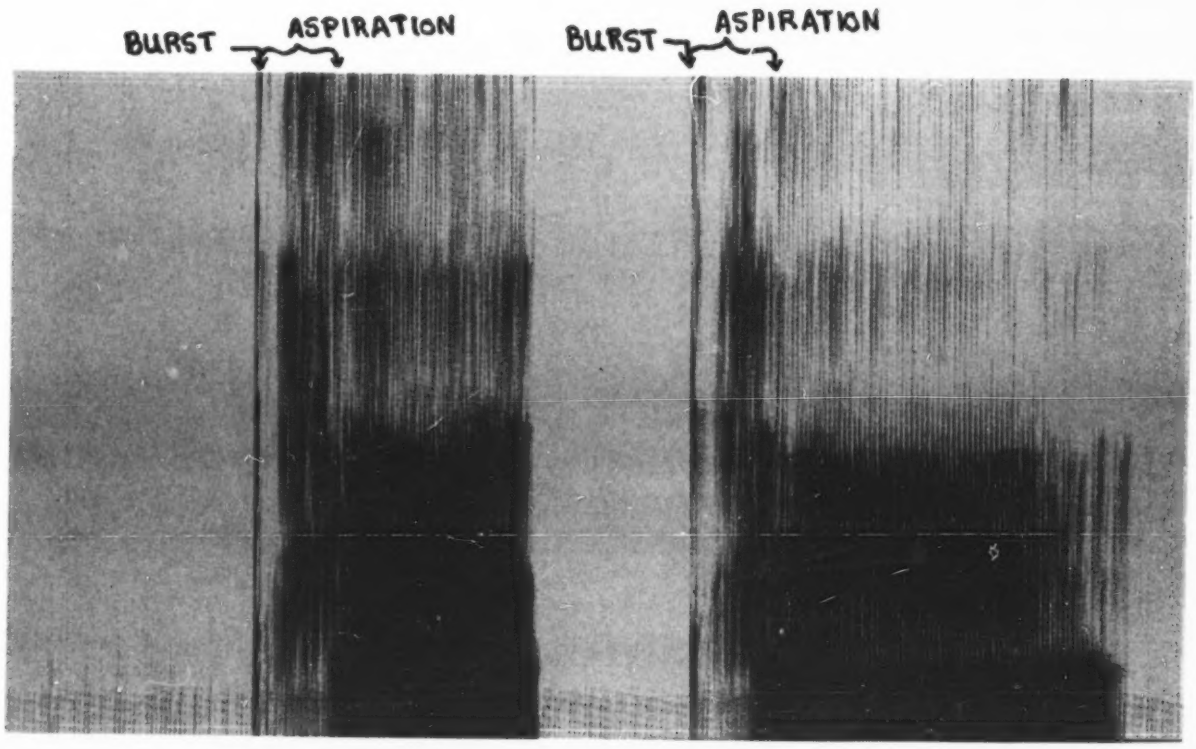




UNVOICED DENTALS

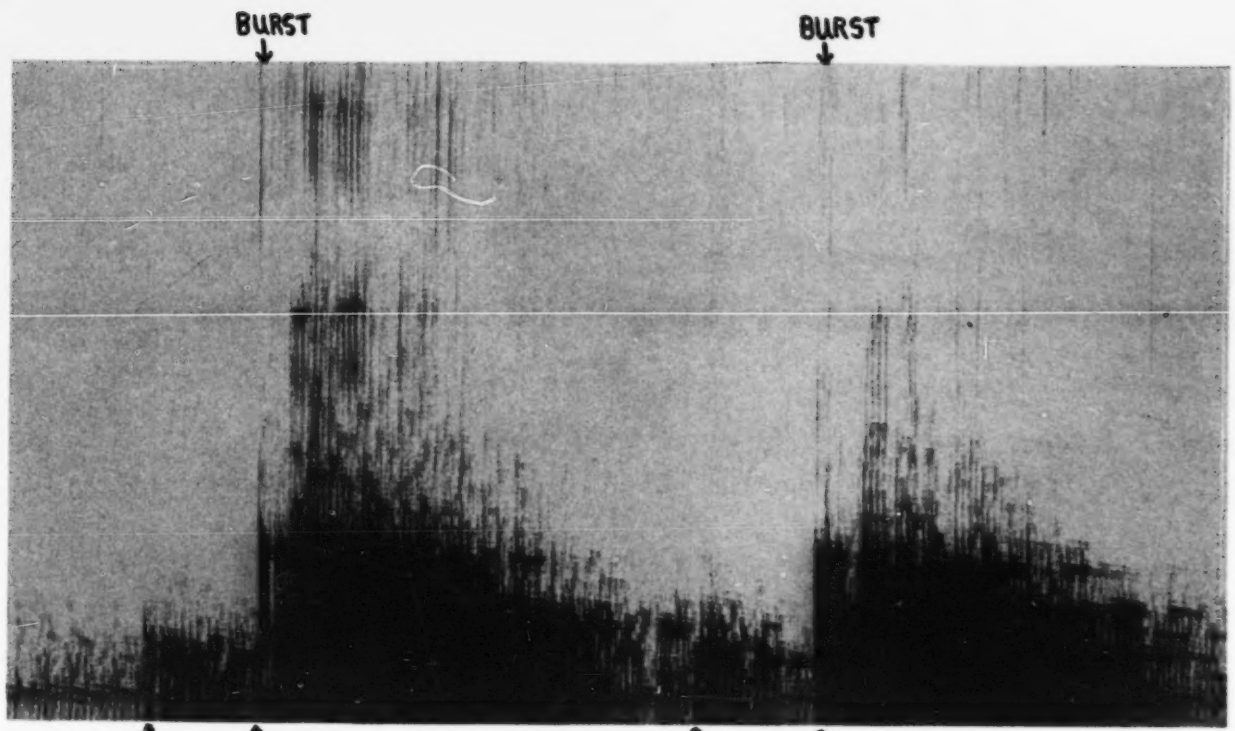


TA TA

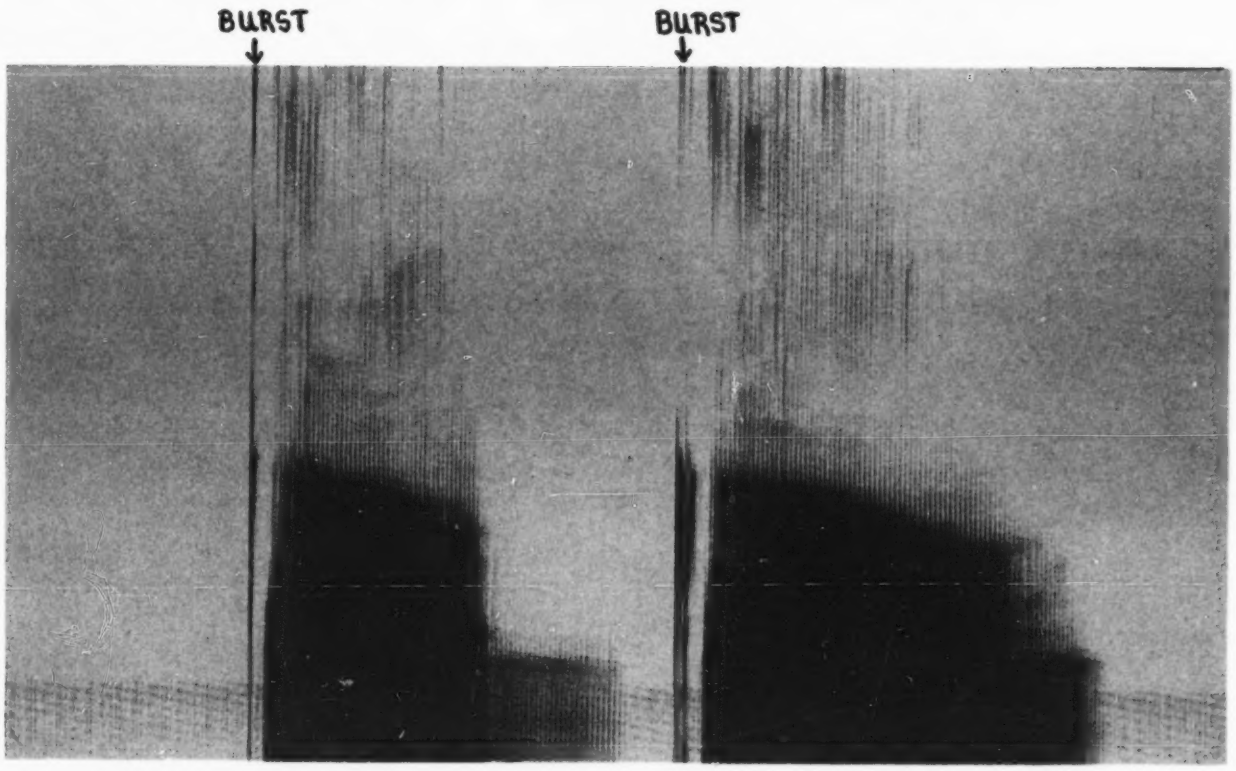


[ta] [ta]

VOICED VELARS

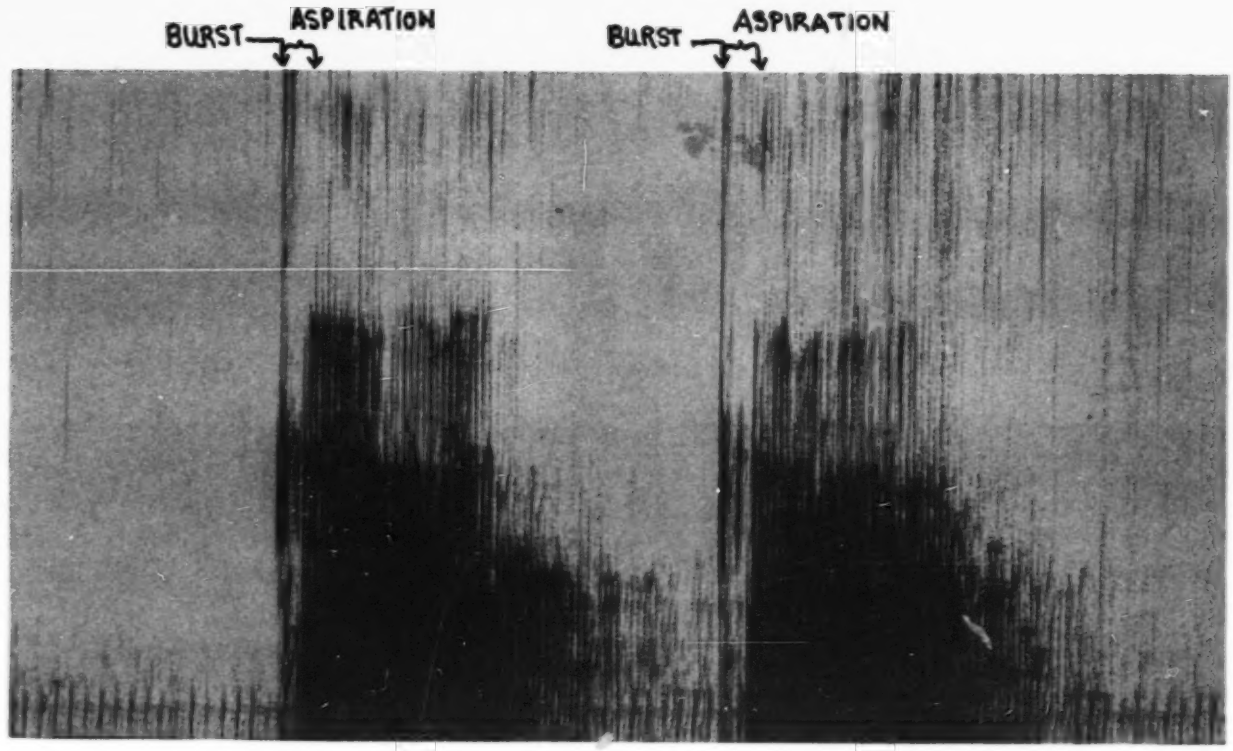


ɡo ɡo

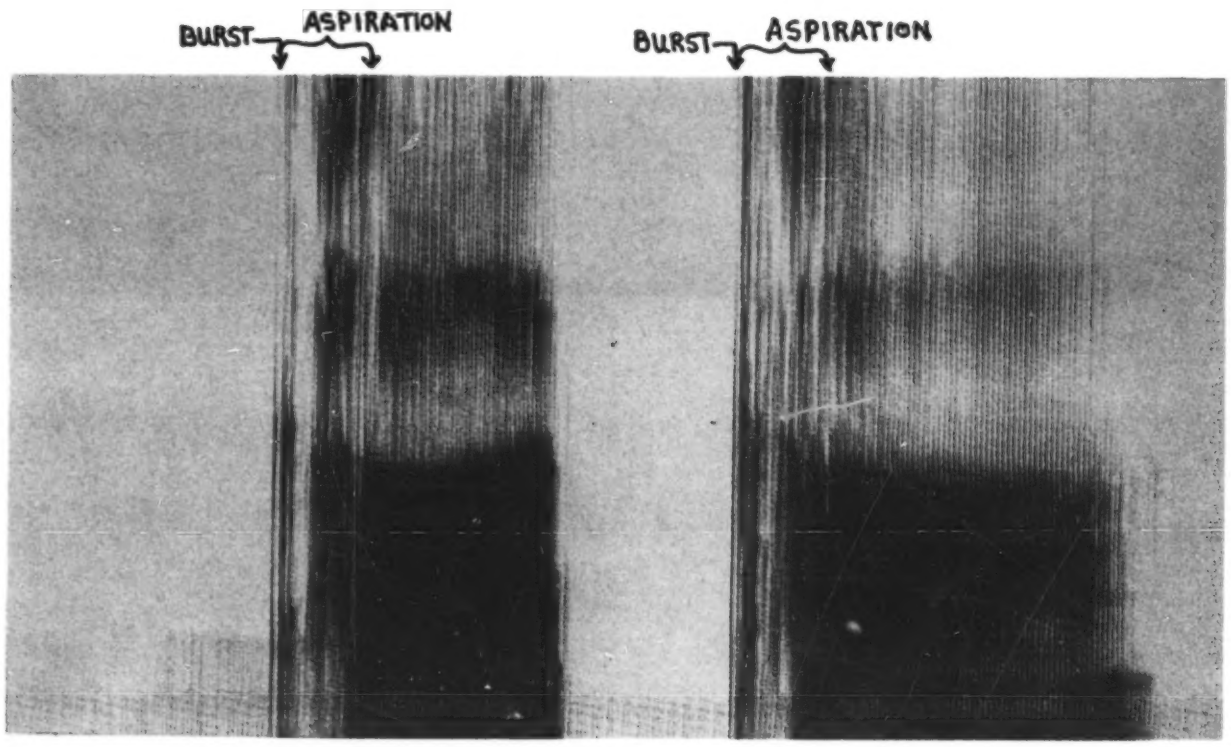


[ɡo] [ɡo]

UNVOICED VELARS



KA KA



[Ka] [Ka]

always longer with the voiceless stops. The exception to this rule is found with bilabials. In Russian, the voiced and voiceless bilabials are not distinguished by relative degrees of aspiration. But in American English they are more clearly distinguished by aspiration than are either dental or velar stops.

There is still another bit of evidence that should be mentioned for its possible relevance to errors with any of the stops, whether they be bilabial, dental, or velar. We refer to evidence regarding the voice bar. For all categories of stops, the voice bar is consistently present in Russian voiced stops and generally absent in American voiced stops. This fact suggests that American listeners may not be accustomed to the voice bar cue and are confused by its presence in Russian syllables.

It becomes apparent then that it is the feature of aspiration which sets the voiceless and voiced bilabial stops characteristically apart from the dento-alveolar and velar ones. We feel, however, that the absence of the burst in both Russian bilabial stops, and the more frequent absence of the voice bar in English bilabials could be additional factors which prevent ready identification by the English-speaking subjects. One word of caution is pertinent here. At this time, the production of only two speakers was subjected to spectrographic analysis; namely, the Russian speaker's who recorded the material for the experiment, and that of one female English speaker's rendition of the above mentioned syllables. We hope to be able to support our tentative conclusions with spectrographic analyses of additional Russian and English speakers in the future.

In List 2, /c/ was the most difficult consonant to be identified, a finding comparable to an earlier one in the previously mentioned report. The least difficult consonant was /č/, as would be expected on the basis of a contrastive English-Russian analysis.

List 3, containing plain vs. palatalized contrasts, yielded data which clearly showed the effect of the following vowel on consonant identification. All but one of the most difficult items contain /i/, the difficulty of which has been shown in our previous reports. The remaining item in this group is difficult both graphemically and phonemically, namely, Л or /l/ vs. /lj/. The grapheme П was chosen frequently for the grapheme Л, and the identification of the plain or palatalized lateral proved difficult for all groups.

For all three lists, the two groups given phonemic transcription did consistently better than the groups using only the Cyrillic alphabet. Specifically, in List 1, Group PA performed best, whereas in Lists 2 and 3 it was $\overline{PA}$. Apparently listening to the alternatives as well as seeing them proved least helpful on the plain vs. palatalized contrasts. This may be due to the wide allophonic range of vowel phonemes after consonants of both series (see Table 6).

Lists 4-6 contained bisyllabic items. The subjects' task was to identify either the stressed or unstressed vowel in one of the syllables, depending on the particular list. In general, the data from these lists suggest that stress patterns linked with vowel identification should have followed consonant identification in monosyllabic CVC patterns, in spite of no anticipated difficulty with the recognition of stress patterns for English speaking subjects. More specifically, in examining the least and most difficult items in these lists, it seems that the subjects tended to

disregard the vowel and listen primarily for stress placement. Thus, among the least difficult items, the stress in the correct alternative usually occurs on the same syllable as in the stimulus, and the reverse tends to be true of the most difficult items. The stress was usually marked on each item on the subject's answer sheet so that perhaps this visual aid also contributed to the above results.

This interpretation is further supported by the fact that no clear pattern of vowel difficulty emerged, in contrast to our previous experiments. (See Suppes et al, 1962b). The phoneme /u/ which had clearly been identified as the least difficult phoneme now appears among the most difficult when in a positionally different stressed syllable than the stimulus vowel. On the other hand, /i/, which previously was without question the most difficult phoneme to identify, was again difficult when the correct answer and the stimulus had different stress patterns, but was among the least difficult when the correct answer and the stimulus had the same stress pattern.

Lists 7 and 8 contained monosyllables of the (C)CVC(C) pattern, the former omitting the problem of neutralization of voicing word finally, the latter including it. The types of mistakes are the same for both lists; that is, the most difficult items usually include final clusters as opposed to single consonants. Also, no palatalized final consonants were found among the easy items. The groups with phonemic transcription did better on both lists than those without, but the difference was more marked on List 7 when the phonemic-graphemic correspondences were closer to a one-to-one relationship, at least in final position. As was mentioned in Section 3, on all the training lists the groups with phonemic transcription were

superior, but on the test they did worse than groups without phonemic transcription. It may be recalled that the test included only Cyrillic orthography and no auditory presentation of the alternatives. As a matter of fact, Group $\overline{PA}$ clearly performed best on the test. The difficulty of /p/ and /b/ discriminations, as well as that of /c/, clearly re-emerged, for all experimental groups. Also, palatalized final consonants presented difficulty in matching, as did items having different stress patterns in the correct answer and stimulus.

Lists 9-11 were based on word matching, so that the previous training on phonemic discriminations could be evaluated in a new context. Generally speaking, items with phonemically very dissimilar alternatives were identified easily by most subjects, whereas those involving minimally contrasting choices proved most difficlt. Most important, these results suggest that training on minimal contrasts did not carry over to longer sequences.

The most striking difference between data from these lists and from previous ones, however, was that the performance of Group $\overline{PA}$ fell much further below that of the other groups (Table 4). For example, in List 9, Group $\overline{PA}$ had only 52% correct responses whereas Group PA had 90%. The combination of phonemic transcription and auditory presentation of the alternatives seems to have a definite bearing on discrimination of longer sequences as opposed to single phoneme matching.

5. EXPERIMENT II

Introduction

This experiment was conducted to determine whether presenting phonemic transcription on only a fraction of the training trials would promote transfer to a test condition in which no transcription was available. As we have already noted, intermittent presentation of transcription might be expected to enhance learning of the Cyrillic more than would presentation on every trial. Of course, the former procedure would also increase the similarity of the training and test conditions, since transcription was never given on test trials. Therefore, the first aim of the present experiment was to test some hypotheses regarding the failure to find facilitating effects of training with transcription in Experiment I. A second purpose was to compare retention and relearning after a 24 hour interval, as against after a two week interval. The experimental task was restricted to the identification of visual representations of whole words in auditorily presented sentences, which was the task in the last three lists of Experiment I. The choice alternatives were not presented auditorily.

6. METHOD

Subjects and experimental design. The subjects were 29 men and 9 women students enrolled in an introductory logic class at Stanford University. Only students who did not know any Russian were eligible for the experiment. Each subject participated in two one-hour sessions and was paid four dollars at the end of the second session.

A 2 x 2 factorial design was used, resulting in four groups of subjects:

		Retention Interval	
		Short	Long
Training with phonemic transcription?	Yes	Group PS	Group PL
	No	Group $\overline{\text{PS}}$	Group $\overline{\text{PL}}$

Subjects were randomly assigned to groups. For Groups $\overline{\text{PS}}$, PS, $\overline{\text{PL}}$, and PL, N = 8, 10, 8, and 12, respectively.

Materials and apparatus. List 9 of Experiment I was used. The other material was the same as in that experiment. To prevent the subject from attending only to the Cyrillic, phonemic transcription never appeared over the correct answer.

Following the fourth list a test was given, consisting of 25 of the training items and 25 new items. As in Experiment I, no phonemic transcription or correct answers were given on the test.

For the PS and PL groups, phonemic transcription was given twice with each item over the first four lists; once in Lists 1 and 2, and once in Lists 3 and 4, making four possible combinations which were counterbalanced across items.

All response alternatives presented were grammatically correct. To reduce the chance that the subject might memorize the correct answer, a different response alternative than the one correct on List 1 was correct on

List 2 for each item. On Lists 3 and 4 these same two correct responses were used. Again the four possible combinations were counterbalanced across items. If an item appeared in the test list, the remaining alternative was used as the correct answer.

Recordings of the instructions and experimental sentences for the subjects were made on an Ampex recorder-reproducer model 601 with an Electrovoice 623 microphone and Ampex Professional Recording Tape of 1 1/2 mil. acetate at 7.5 inches per second. The person recording the material was the same native Russian speaker who recorded for our pervious experiments. The instructions were recorded by a native American speaker. The recordings were played using the Ampex 601 tape recorder-reproducer and an Ampex 620 amplifier loudspeaker. The play-back locale was a room 15 ft. x 25 ft..

The total time for each item was approximately 7 sec.. The length of each list was: List 1, 6 min. 50 sec., List 2, 7 min. 5 sec., List 3, 7 min. 20 sec., List 4, 7 min. 25 sec., and List 5, 8 min. 4 sec.. Pauses were inserted on the tapes to allow the subjects time to turn pages and to move from list to list. Brief instructions, similar to those in Experiment I, were given at the beginning of the experiment and between Lists 4 and 5. The total time for an experimental session was approximately 50 min..

Procedure. The subjects were seated at a large table and each had Lists 1 through 5 and a blotter in front of him. The instructions were like those used in List 9 of Experiment I (Appendix B). They were played and any questions the subjects had were answered by paraphrasing the relevant portion of the instruction. Following this, Lists 1 through 4 were presented. The subject circled one of the alternatives and then moved the blotter down to compare his choice with the correct answer. He then proceeded to the next item in the same fashion. At the end of List 4 the second set of instructions was given and then the test list was presented.

The second session for all subjects was identical in procedure to session 1. For $\overline{PS}$ and PS groups the second session was held 24 hours after the first session and for the $\overline{PL}$ and PL groups, the interval between sessions was 2 weeks.

7. RESULTS OF EXPERIMENT II

The proportion correct for each group on each list is shown in Figure 4. The amount of improvement from list to list is remarkedly slight. To illustrate the effect of the retention interval, the data of the two groups with the same retention interval were pooled (Figure 5). Similarly, the effect of transcription vs. no transcription was exhibited by pooling over retention intervals (Figure 6). We shall briefly comment on these figures before turning to the major analysis, which was an analysis of variance in total number of correct responses on the Day 2 test.

Insert Figures 4, 5, and 6 about here

According to Figure 5, the Day 1 proportions correct were about .05 higher for the long interval group than for the short interval group. On Day 2 the discrepancy dropped to about .02. Since the retention interval variable did not enter until after Day 1, any inequality in Day 1 proportions correct would imply that extraneous factors were operating (e.g., bias in the selection of subjects). However, an analysis of variance in the total number of correct responses on the Day 1 test failed to yield any significant inter-group differences (for the retention interval $F = 3.70$, which was not significant).

Figure 6 displays the comparison between the groups with and without phonemic transcription training. As in Experiment I, the training led to a higher proportion correct when the transcription was present, (Lists 1-4 and 6-9). The advantage did not transfer to a test when transcription was no longer available. That is, on the Day 1 test, the proportions correct were .57 for the no-transcription group and .58 for the transcription group, respectively. The corresponding proportions on the Day 2 test were .54 and .58.

Two analyses of variance were computed from the data of each day, once using the total correct on the training lists and again using the total correct on the test. We were concerned chiefly with the relatively long-lasting effects -- those discernible from the Day 2 data.

Table 7 shows the results of an analysis of variance in total number of correct responses on the Day 2 training lists. The F ratio between the

---

Insert Table 7 about here

---

transcription and no-transcription groups is highly significant ($F = 11.15$, $P < .01$). We have already mentioned that this finding agrees with Experiment I. Since the effect of the interval between sessions is not significant, any initial retention differences must have been cancelled by relearning.

An analysis of variance in the total number of correct responses on the Day 2 test (List 10) appears in Table 8. Neither the main effects nor

---

Insert Table 8 about here

---

the interaction was significant. The same result obtained for the Day 1 test. Evidently, then, prior training with phonemic transcription did not affect performance on a test in which no transcription was presented. In view of the above-mentioned failure to find a significant effect of the retention interval on the Day 2 training lists, it is not surprising that no difference was found on the Day 2 test.

Since the tests contained new items as well as items which had been presented during the training series, it was possible to compare the groups separately on the two item types. Table 9 gives the proportion correct on

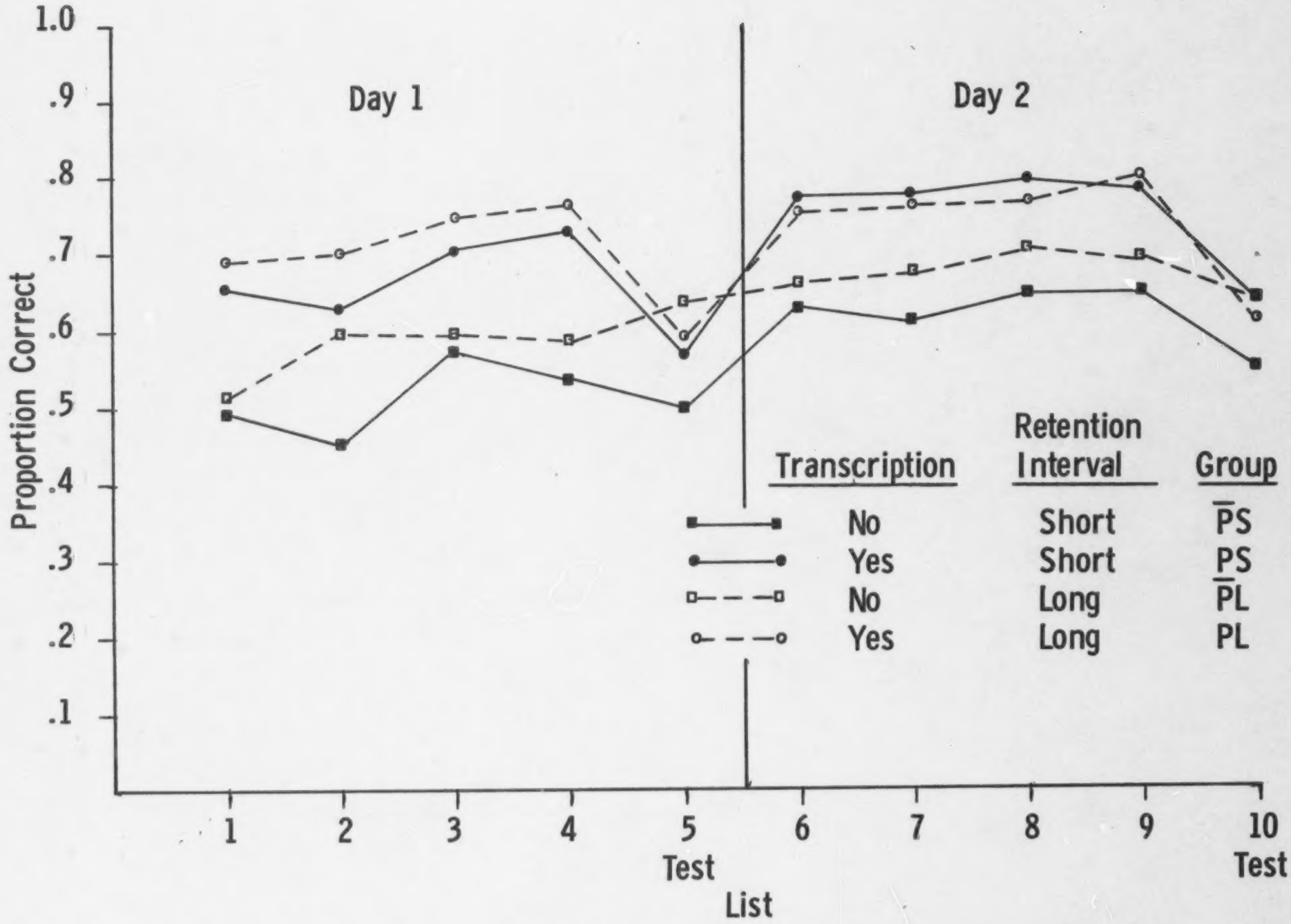


Figure 4. Proportion correct for each group on each list.

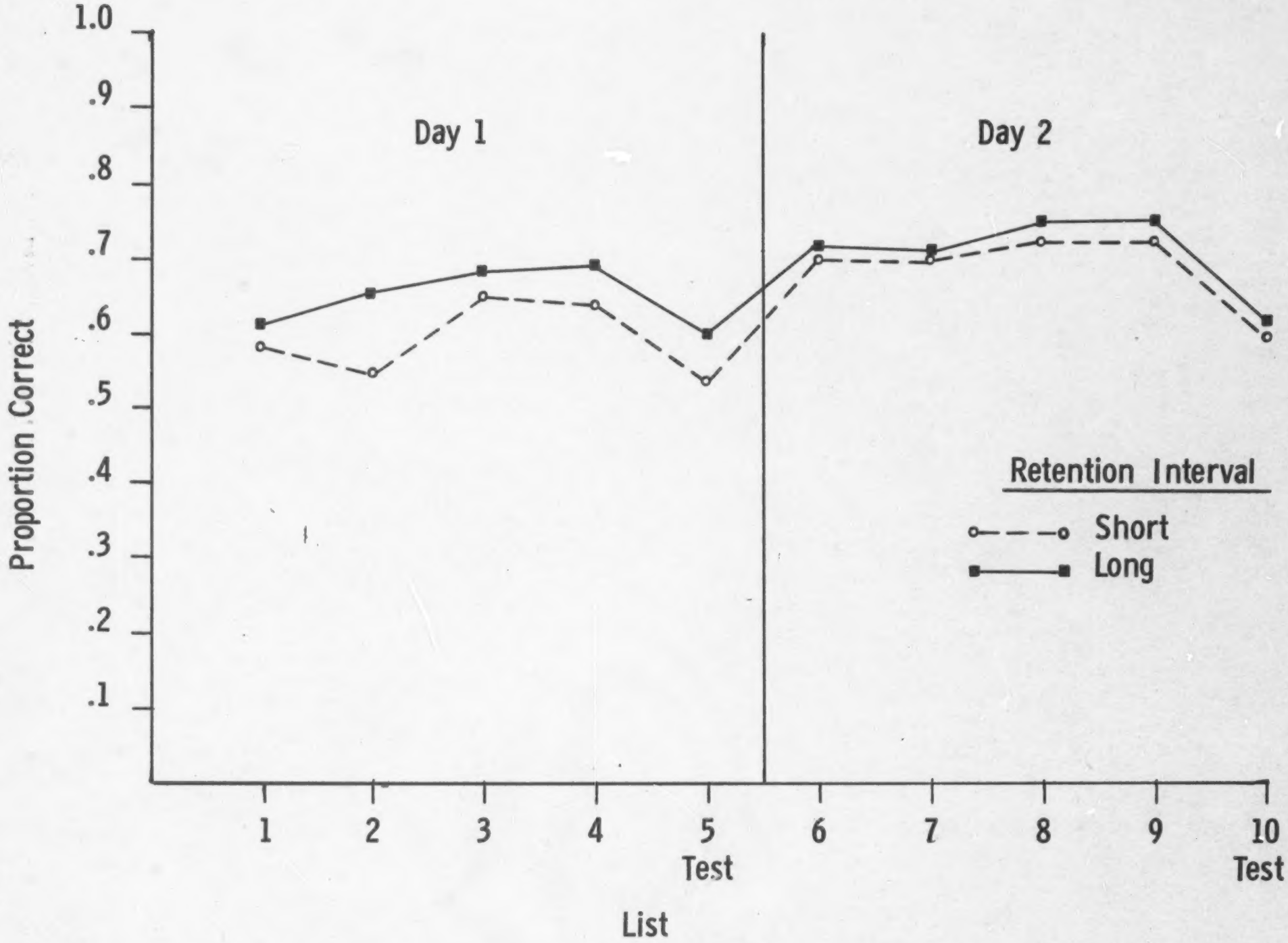


Figure 5. Proportion correct for short and long retention groups on each list.

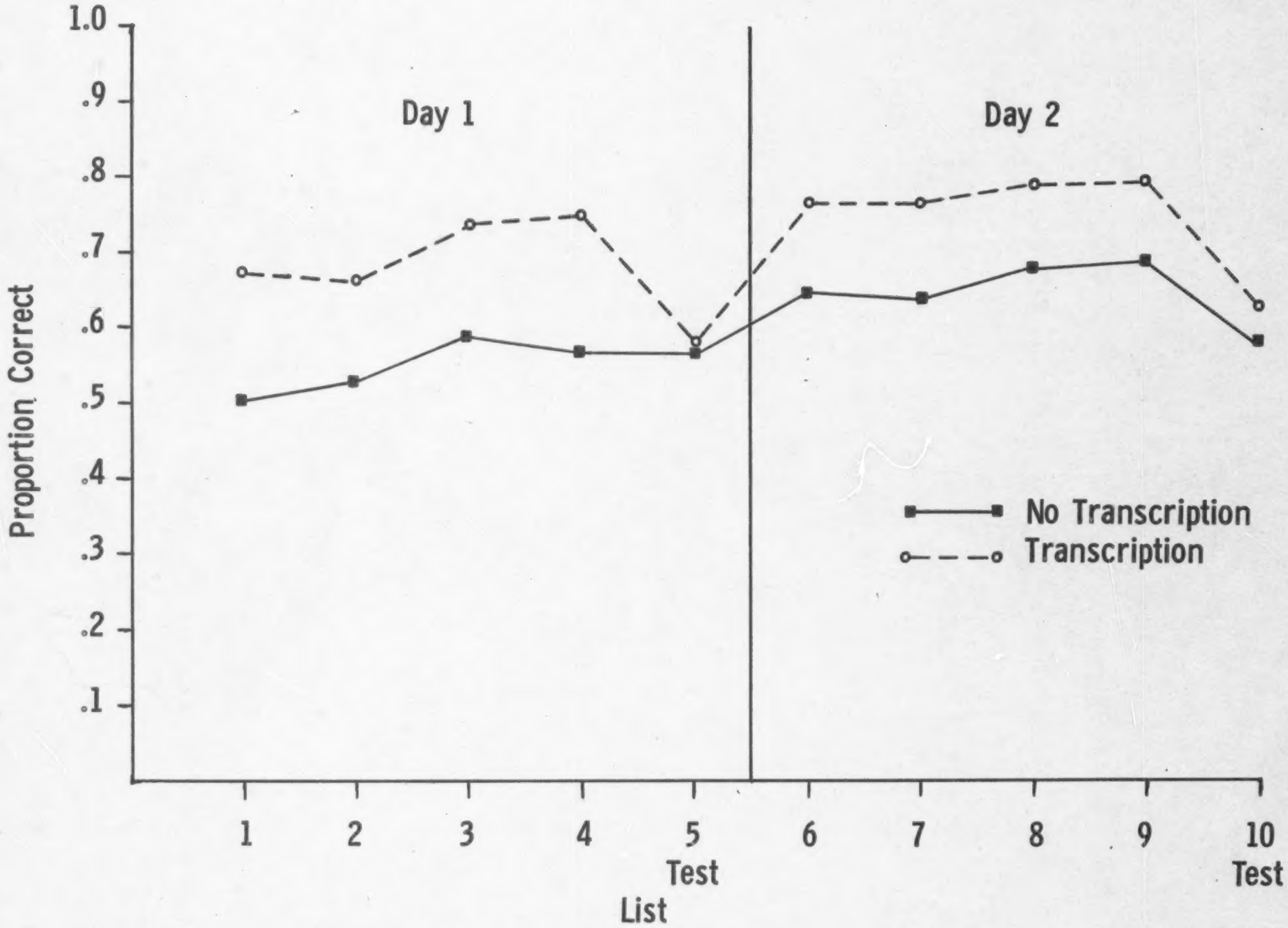


Figure 6. Proportion correct for groups with and without transcription on each list.

Table 7

Results of Analysis of Variance of Total Correct in Lists 6-9

Source	Mean Square	df.	F
Retention Interval	77.55	1	0.17
Experimental Condition	5051.47	1	11.15*
Interaction	417.14	1	0.92
Between	1848.72	3	
Within	453.03	34	

* $P < .01$

Table 8

Results of Analysis of Variance of Total Correct in List 10

Source	Mean Square	df.	F
Retention Interval	23.42	1	0.53
Experimental Condition	26.56	1	0.61
Interaction	71.09	1	1.62
Between	40.35	3	
Within	43.82	34	

Insert Table 9 about here

---

the old and on the new items for the group with and the group without prior transcription training. On both tests, the proportion correct on old items was .01 higher for the group which had transcription training than for the group without this training. The difference reached .06 for the new items on the Day 2 test. However, none of these differences were statistically significant.

Although we were primarily interested in the relearning measure of retention, some indirect evidence is available regarding recognition memory from the end of the first session to the start of the second. The probability of a correct response on List 6, conditional on a correct response on List 4, was taken as a rough index of retention. Of course, the contaminating factors are the responses on the first test and relearning during List 6. For the groups with transcription, this conditional probability was computed once over the items having transcription present on both Lists 4 and 6. The same items were included in the computations for the groups without transcription. Table 10 presents the results of this analysis. The only noteworthy difference

---

Insert Table 10 about here

---

between the short and long retention interval groups involved the items for which transcription was present on both Lists 4 and 6. Here the proportion correct was .95 for the short interval group against .86 for the long interval group. This finding suggests that a short retention interval may permit more recall of the transcription, but the effect is rather mild.

A final tangential analysis was performed to determine whether certain sequential dependencies were present in the training list data of the group with transcription. The question was whether the probability of a correct response depended on whether transcription had been available on the preceding presentation of the same item (i.e., the presentation in the preceding list). If we again denote the presence of phonemic transcription on trial n by P_n and the absence by $\overline{P}_n$, the four combinations $P_{n-1}P_n$, $P_{n-1}\overline{P}_n$, $\overline{P}_{n-1}P_n$, and $\overline{P}_{n-1}\overline{P}_n$ of successive presentations of the same item occurred equally often. Then two comparisons of conditional probabilities are of interest. First, we consider the $P_{n-1}P_n$ and $\overline{P}_{n-1}P_n$ items, asking whether the probability of a correct response on trial n for items with transcription depends on whether transcription were present on the previous presentation of that item (rows 1 and 2 of Table 11). Likewise, we compare the items without transcription of the trial in question -- the $P_{n-1}\overline{P}_n$ and $\overline{P}_{n-1}\overline{P}_n$ items, (rows 3 and 4 of Table 11). All the comparisons yielded negligible differences, so no further statistical tests

---

Insert Table 11 about here

---

seem required to establish that there was no prominent sequential dependency of this type.

Let us briefly comment on the amount of learning found and recapitulate our conclusions from these two recognition experiments. The following remarks concerning the effect of the retention interval are made with reference to Figure 6 of Experiment II. Nearly all the improvement occurred on List 1, since the proportion correct on this list was .60, in contrast to the .33 expected by chance. Why learning did not continue is not clear.

Table 9

Proportion Correct for Old and New Items
in Test Lists

		Proportion Correct	
		Group	
List 5 (Day 1)	Type	No transcription	Transcription
	Old	.60	.61
	New	.53	.55
List 10 (Day 2)			
	Old	.64	.65
	New	.54	.60

Table 10

Proportion Correct on the First Training List on Day 2 (List 6), Conditional on a Correct Response on the Last Training List on Day 1 (List 4)

	Retention Interval	
	Short	Long
No transcription group	.67	.70
Transcription group		
All items $\bar{P}$ on both lists	.63	.59
All items P on both lists	.95	.86

Table 11

Proportion of Correct Responses to Items Classified According to Whether Phonemic Transcription was Available (P) or not Available ($\overline{P}$) on the Preceding Trial and on the Present Trial

Preceding Trial	Present Trial	Proportion Correct on Present Trial	
		Day 1	Day 2
P	P	.76	.79
$\overline{P}$	P	.70	.78
P	$\overline{P}$	.72	.81
$\overline{P}$	$\overline{P}$	.75	.76

The major finding was that the availability of phonemic transcription during the learning of Russian auditory - orthographic associations does not appear to improve performance on a test when the transcription is no longer present. Nor did learning with auditory presentation of the choice alternatives lead to higher scores on the test. Finally, the comparison of short vs. long inter-session intervals in Experiment II revealed that retention differences, as measured from Day 2 performance, were negligible. It seems plausible that the lack of continued improvement from list to list at least partly accounts for the failure to find a persistent effect of the retention interval. If a higher level of learning had been reached on Day 1, more forgetting during the interval might have appeared. In both experiments, our claim that phonemic transcription is of very limited value in facilitating learning has only been tested and affirmed under the condition of partial, not complete, learning.

8. REFERENCES

Reeds, J.A., and Wang, W. S. Y., The perception of stops after "s". Phonetica, 1961, 6, 78-81.

Suppes, P., Crothers, E., and Weir, R. Some quantitative studies of Russian consonant phoneme discrimination. Tech. Rep. No. 49, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1962a.

Suppes, P., Crothers, E., and Weir, R. Application of mathematical learning theory and linguistic analysis to vowel phoneme matching in Russian words. Tech. Rep. No. 51, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1962b.

APPENDIX A

Sample Items from Each List

List				
1.	РО-БО	РО-БО	ЛО-БО	БО-БО
2.	ЦА-СА	ЩА-СА	СА-СА	ЦА-СА
3.	НЫ-НИ	ХИ-НИ	НЫ-НИ	НИ-НИ
4.	БО́НА	ПЕРО́	МЯ́ГИ	ГОРЮ́
5.	МЮ́РЯ	ЛУ́ПЫ	ДЫ́НЯ	МОРЯ́
6.	ГА́ЛЕ	ДУ́НЕ	ЦУ́КА	МОСЬЁ
7.	ЛЕНЧ	СЬЕСТЬ	ВОЛК	ХАРЧ
8.	ВЕТЬ	ВЁС	ГУТЬ	ПЫШ
9.	ГОРЫ́	ЛЕ́СА	ГОРЫ́	ПРУДА́
10.	КОМ	КОМ	ПОП	СЭР
11.	РУКЁ	РУКЁ	НЕ́БУ	ПИ́ЩЕ

APPENDIX B

Experiment IX Instructions

List 4, Regular Stressed Vowels. You are going to hear a list of Russian words. Each of the key Russian words has two syllables, one of which is more loudly stressed than the other. The stressed syllable will have one of five vowel sounds, approximately like the /a/ in father, the /e/ in Betty, the /o/ in foresight, the /u/ in ruler, the /i/ in milky or pizza. Listen to which vowel has the loud stress in the key word. Decide which of the three alternatives on your answer sheet has the same vowel, with the same loud stress, as the key word. Circle this answer. Then see if you were right by looking at the word in the first column of the next line. Then listen to the next key word and continue as before.

List 5, Regular Unstressed Vowels. You are going to hear a list of Russian two syllable words. Like the words in the preceding session, one syllable is stressed and the other is unstressed, or quieter. This time, you are to listen for the unstressed vowel in the key word. The unstressed syllable will have one of three vowel sounds, approximately like the /a/ in father, the /u/ in ruler, the /i/ in milk or pizza, and will be softer than the other syllable. After you have decided which vowel in the key word is unstressed, examine the three alternatives on the sheet in front of you. Circle the word that has the same unstressed vowel as the key word. After you have answered, look at the first column of the line below to see if your answer was correct. Then listen for the unstressed vowel in the next key word.

List 6, Irregular Unstressed Vowels. As with List 5 in the preceding session, your task is to identify the unstressed vowel in the key word and circle the answer which contains the same unstressed vowel on your answer sheet. It will help you to know that in Russian the letter than looks like O sounds like this when unstressed: /a/ or /ə/, and that the letter that looks like E when it is unstressed sounds like this: /i/. Remember, each time you listen, to check your answer for correctness.

List 7, Regular CVC. You are going to hear a list of one-syllable Russian words. Each of the key words ends in one or more consonants. Listen for the very last consonant sound, and decide which one ends in the same consonant sound as the key word. Circle your answer, then look at the first column of the line below to see if it is correct. Then listen for the last consonant sound of the next key word.

List 8, Irregular CVC. As with List 7, in the preceding session, your task is to identify the final consonant sound of the key word. Listen for the very last consonant sound. Then decide which alternative ends in the same consonant sound as the key word. It will help you to know that there are six Russian letters which have a differenent pronunciation from the usual one when they are at the end of a word.

Д	sounds	more	like	a	t	than	a	d
Б	"	"	"	"	p	"	"	b
В	"	"	"	"	f	"	"	v
Г	"	"	"	"	k	"	"	g
Ж	"	"	"	"	sh	"	"	zh
З	"	"	"	"	s	"	"	z

Your task is still to decide which alternative has the same final sound as the key word.

Instructions for Test based on Lists 1-8.

You are going to have a test on three kinds of items selected from the lists you have been working with up until now. One kind of item is the consonant-vowel pair in which, for example, you hear this: [mi-mji] and then select the appropriate spelling [МЫ-МИ] from the three alternatives on your answer sheet. Another example of this kind of pair is this: [ta-da] which would correspond to the appropriate spelling [ТА-ДА] on your sheet. The second kind of item is the one-syllable word ending in a consonant which you must identify. For instance if you hear this: [porsh] [ПОРШ] you pick the alternative which also ends in the sound [sh], for instance [gyelysh] [ГЕЛШ]. The third kind of item is the two-syllable word in which you listen for the unstressed vowel. For instance, if you hear this: [nári] [НАРЫ] the unstressed vowel is [i] [Ы] and the correct alternative to pick might be [lúri] [ЛУ́РЫ] which also ends in unstressed [i] [Ы]. Another example of this third kind of item might be this: [gúpa], [ГУ́ПА] with an unstressed [a] [A], and the right answer is [shalí] [ШАЛЫ] which also has an unstressed [a] [A]. Only these three kinds of items are included, the consonant-vowel pair, the one-syllable word whose final consonant you must identify, and the two-syllable word whose unstressed vowel is what you listen for.

Since this is a test, you will of course not be given the right answer. Thus there are only three columns of words on your answer sheet, which are the three alternatives for each item. For some of you, there will be other differences between the test procedure and the experimental procedure which you have been following. On the test, you will see only the Russian spellings, and you will look at the alternatives, not hear them.

Lists 9-10, 2-word sentences. You are going to hear a list of short Russian sentences, consisting of two words of various lengths. From the three alternatives on the sheet before you, decide which word was one of the 2 words you heard in the sentence. Circle the word which corresponds to one of the two words you heard in the two-word key sentence. Look at the first column of the line below to see if your answer is correct. Then listen to the next key sentence, and again, decide which of the three alternative words was the one that occurred in the sentence. Continue to the end in this manner.

List 11 has longer sentences.

You are going to hear a list of Russian sentences of various lengths. From the three alternatives on the sheet before you, decide which word was the one you heard in the sentence. Look at the first column of the line below to see if your answer is correct. Then listen to the next key sentence, and again, decide which of the three alternative words was the one that occurred in the sentence. Continue to the end in this manner.